GenCore version 6.1.x
Copyright (c) 1993 - 2002 Compugen Ltd.

Mpi*tab - protein search, using sw model

Run on: February 5, 2003, 12:44:20 ; Search time 7.2087 Seconds
(without alignments)
200.219 Million cell updates/sec

Title: US-09-423-100-1
Perfect score: 260
Sequence: 1 MPPIIELSI[illegible]

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262774 seqs, 264,[illegible] residues

Total number of hits satisfying chosen parameters: 262774

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/ina/Issued_Patents_AA_pep:*
2: /cgn2_6/ptodata/1/ina/Issued_Patents_AB_pep:*
/cgn2_6/ptodata/1/ina/Issued_Patents_AC_pep:*
4: /cgn2_6/ptodata/1/ina/Issued_Patents_AD_pep:*
5: /cgn2_6/ptodata/1/ina/Issued_Patents_AE_pep:*
6: /cgn2_6/ptodata/1/ina/Issued_Patents_AF_pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	260	100.0	1[illegible]	4	US-[illegible] Sequence 1, Appli
2	259	99.6	1[illegible]	4	US-[illegible] Sequence 5, Appli
3	257	98.8	1[illegible]	4	US-[illegible] Sequence 4, Appli
4	257	98.8	1[illegible]	4	US-[illegible] Sequence 4, Appli
5	256	98.5	2[illegible]	4	US-[illegible] Sequence 10, App
6	256	98.5	2[illegible]	4	US-[illegible] Sequence 1, App
7	256	98.5	2[illegible]	4	US-[illegible] Sequence 10, App
8	256	98.5	2[illegible]	4	US-[illegible] Sequence 11, App
9	256	98.5	2[illegible]	4	US-[illegible] Sequence 1, App
10	256	98.5	241	4	US-[illegible] Sequence 2, App
11	256	98.5	274	4	US-[illegible] Sequence 71, App
12	256	98.5	[illegible]	4	US-[illegible] Sequence 73, App
13	244	93.8	191	4	US-[illegible] Sequence 1, Appl
14	243	93.5	217	4	US-[illegible] Sequence 4, Appl
15	241	92.7	217	1	US-[illegible] Sequence 51, App
16	241	92.7	217	2	US-[illegible] Sequence 4, App
17	238	91.5	191	4	US-[illegible] Sequence 16, App
18	238	91.5	191	4	US-[illegible] Sequence 16, Appl
19	238	91.5	[illegible]	4	US-[illegible] Sequence 2, App
20	237	91.2	191	1	US-[illegible] Sequence 24, App
21	237	91.2	191	1	US-[illegible] Sequence 24, App
22	[illegible]	[illegible]	[illegible]	[illegible]	[illegible] Sequence 20, App
23	236	90.8	191	1	US-[illegible] Sequence 8, Appli
24	234	90.0	[illegible]	4	US-[illegible] Sequence 1, App
25	234	90.0	[illegible]	4	US-[illegible] Sequence 1, App
26	233	89.6	191	1	US-[illegible] Sequence 4, Appli

ALIGNMENTS

RESULT 1
[illegible]
; Sequence 1, Application [illegible]
; Patent No. [illegible]
; GENERAL INFORMATION:
; APPLICANT: [illegible]
; APPLICANT: [illegible]
; TITLE OF INVENTION: DNA [illegible]
; NUMBER OF SEQUENCES: [illegible]
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: [illegible]
; STREET: [illegible]
; CITY: [illegible]
; STATE: [illegible]
; COUNTRY: USA
; ZIP: [illegible]
; COMPUTER READABLE FORM:
; MEDIUM TYPE: [illegible]
; COMPUTER: IBM [illegible]
; OPERATING SYSTEM: [illegible]
; SOFTWARE: [illegible]
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: [illegible]
; FILING DATE: [illegible]
; CLASSIFICATION: [illegible]
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: [illegible]
; FILING DATE: [illegible]
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: [illegible]
; FILING DATE: [illegible]
; ATTORNEY/AGENT INFORMATION:
; NAME: [illegible]
; REGISTRATION NUMBER: [illegible]
; REFERENCE/DOCKET NUMBER: [illegible]
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: [illegible]
; TELEFAX: [illegible]
; TELEX: [illegible]
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: [illegible]
; TYPE: [illegible]

; ... (text)
US-08-xxx-xxx-1

Query Match 100.0%; Score 260; DB 1; Length 192;
Best Local Similarity 100.0%; Pred. No. 1.5e-17;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEPITPGSRLFNAMCAARHSLIAFEDYGQEPDAYTRKLKKYSPTANPLK 48
||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MEPITPGSRLFNAMCAARHSLIAFEDYGQEPDAYTRKLKKYSPTANPLK 48

RESULT 2
US-09-284-878-5
; Sequence 5, Application US/09284878
; Patent No. 6541575
; GENERAL INFORMATION:
; APPLICANT: Navarro, Martin Gutierrez
; APPLICANT: Saldana, Hugo Barrera
; APPLICANT: Saldana, Jose Maria Viedet
; TITLE OF INVENTION: Genetically Modified Methylotrophic P. pastoris Yeast for the
; TITLE OF INVENTION: Production and Secretion of the Human Growth Hormone
; FILE REFERENCE: 1829.001Dom0
; CURRENT APPLICATION NUMBER: US/09/284,878
; CURRENT FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: PCT/MX97/00003
; PRIOR FILING DATE: 1997-10-24
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 5
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-284-878-5

Query Match 99.1%; Score 258; DB 1; Length 192;
Best Local Similarity 100.0%; Pred. No. 1.5e-17;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEPITPGSRLFNAMCAARHSLIAFEDYGQEPDAYTRKLKKYSPTANPLK 48
||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MEPITPGSRLFNAMCAARHSLIAFEDYGQEPDAYTRKLKKYSPTANPLK 48

RESULT 3
US-08-383-621-4
; Sequence 4, Application US/08383621
; Patent No. 5951972
; GENERAL INFORMATION:
; APPLICANT: Daley, Michael J.
; APPLICANT: Buckwalter, Brian L.
; APPLICANT: Eady, Susan M.
; APPLICANT: Shieh, Hong-Ming
; APPLICANT: Bolden, Peter
; APPLICANT: Sodum, Adit R.
; TITLE OF INVENTION: Stabilization of Somatotropins And Other
; TITLE OF INVENTION: Proteins By Modification Of Cysteine Residues
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. Patricia J. Verdesca
; STREET: 1767 West Main Street, P.O. Box 5
; CITY: Stamford
; STATE: Connecticut
; COUNTRY: U.S.A.
; ZIP: 06904-0005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/383,621
; FILING DATE: 06-FEB-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/xxx,xxx
; FILING DATE: 25-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Isovdos, Estelle J.
; REGISTRATION NUMBER: 31,119
; REFERENCE/DOCKET NUMBER: [illegible]
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 203-321-2116
; TELEFAX: 203 321 2971
; TELEX: 203 321-3434-4600
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 194 amino acids
; TYPE: amino acid
; STRANDEDNESS: linear
; MOLECULE TYPE: protein
US-08-383-621-4

Query Match 98.1%; Score 255; DB 1; Length 194;
Best Local Similarity 100.0%; Pred. No. 1.5e-17;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EPITPGSRLFNAMCAARHSLIAFEDYGQEPDAYTRKLKKYSPTANPLK 49
||||||||||||||||||||||||||||||||||||||||||||||||
Db 4 EPITPGSRLFNAMCAARHSLIAFEDYGQEPDAYTRKLKKYSPTANPLK 51

RESULT 4
US-08-459-906-4
; Sequence 4, Application US/08459906
; Patent No. 6010999
; GENERAL INFORMATION:
; APPLICANT: Daley, Michael J.
; APPLICANT: Buckwalter, Brian L.
; APPLICANT: Eady, Susan M.
; APPLICANT: Shieh, Hong-Ming
; APPLICANT: Bolden, Peter
; APPLICANT: Sodum, Adit R.
; TITLE OF INVENTION: Stabilization of Somatotropins and Other
; TITLE OF INVENTION: Proteins By Modification of Cysteine Residues
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: American Cyanamid Company
; STREET: One Cyanamid Plaza
; CITY: Wayne
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 07470-8426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,906
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Barry L.
; REGISTRATION NUMBER: 34,758
; REFERENCE/DOCKET NUMBER: 51,274-CON
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 831 4247
; TELEFAX: 201 831 [illegible]
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 194 amino acids
; TYPE: amino acid
; STRANDEDNESS: linear
; MOLECULE TYPE: protein
US-08-459-906-4

```
;         STATE: Texas
;         COUNTRY: USA
;         ZIP: [illegible]
;       COMPUTER READABLE FORM:
;         MEDIUM TYPE: Floppy disk
;         COMPUTER: IBM PC compatible
;         OPERATING SYSTEM: PC-DOS/MS-DOS
;         SOFTWARE: PatentIn Release #1.0, Version #1.3
;       CURRENT APPLICATION DATA:
;         APPLICATION NUMBER: US/08/785,[illegible]
;         FILING DATE: Concurrently Herewith
;         CLASSIFICATION: 435
;       PRIOR APPLICATION DATA:
;         APPLICATION NUMBER: US 08/589,028
;         FILING DATE: [illegible]-JAN-[illegible]
;       ATTORNEY/AGENT INFORMATION:
;         NAME: Highlander, Steven L.
;         REGISTRATION NUMBER: [illegible]
;         REFERENCE/DOCKET NUMBER: [illegible]
;       TELECOMMUNICATION INFORMATION:
;         TELEPHONE: 512/[illegible]
;         TELEFAX: 512/474-[illegible]
;   INFORMATION FOR SEQ ID NO: 10:
;       SEQUENCE CHARACTERISTICS:
;         LENGTH: 217 amino acids
;         TYPE: amino acid
;         STRANDEDNESS:
;         TOPOLOGY: linear
US-08-785-271-10

  Query Match             98.1%;  Score 255;  DB 1;  Length 217;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-29;
  Matches   48;  Conservative    [illegible];  Mismatches    [illegible];  Indels    [illegible];  Gaps    [illegible];

Qy          2 EPEIPLSRLEPNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db         27 EPEIPLSRLEPNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 74


RESULT 8
US-08-759-628-11
; Sequence 11, Application US/08759628
; Patent No. 6225[illegible]
;   GENERAL INFORMATION:
;       APPLICANT: Altmann, Scott W.
;       APPLICANT: Rock, Fernando L.
;       APPLICANT: Bazan, J. Fernando
;       APPLICANT: Kastelein, Robert A.
;       TITLE OF INVENTION: [illegible]
;       NUMBER OF SEQUENCES: 11
;       CORRESPONDENCE ADDRESS:
;         ADDRESSEE: DNAX Research Institute
;         STREET: [illegible] California Avenue
;         CITY: Palo Alto
;         STATE: California
;         COUNTRY: USA
;         ZIP: 94304-1104
;       COMPUTER READABLE FORM:
;         MEDIUM TYPE: Floppy disk
;         COMPUTER: IBM PC compatible
;         OPERATING SYSTEM: PC-DOS/MS-DOS
;         SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;         APPLICATION NUMBER: US/08/759,628
;         FILING DATE: 05-DEC-[illegible]
;         CLASSIFICATION: 435
;       PRIOR APPLICATION DATA:
;         APPLICATION NUMBER: US 60/008,574
;         FILING DATE: [illegible]-DEC-[illegible]
;       ATTORNEY/AGENT INFORMATION:
;         NAME: Ching, Edwin P.
;         REGISTRATION NUMBER: [illegible]
;         REFERENCE/DOCKET NUMBER: [illegible]
;       TELECOMMUNICATION INFORMATION:
;         TELEPHONE: 415-[illegible]
;         TELEFAX: 415-[illegible]
;   INFORMATION FOR SEQ ID NO: 11:
;       SEQUENCE CHARACTERISTICS:
;         LENGTH: 217 amino acids
;         TYPE: amino acid
;         STRANDEDNESS: single
;         TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       FEATURE:
;         NAME/KEY: Peptide
;         LOCATION: [illegible]
;       FEATURE:
;         NAME/KEY: Peptide
;         LOCATION: 94..115
;       FEATURE:
;         NAME/KEY: Peptide
;         LOCATION: [illegible]
;       FEATURE:
;         NAME/KEY: Peptide
;         LOCATION: [illegible]
;         OTHER INFORMATION: /note= "The peptides above are
;         OTHER INFORMATION: [illegible]"
US-08-759-628-11

  Query Match             98.1%;  Score 255;  DB 1;  Length 217;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-29;
  Matches   48;  Conservative    [illegible];  Mismatches    [illegible];  Indels    [illegible];  Gaps    [illegible];

Qy          2 EPEIPLSRLEPNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db         27 EPEIPLSRLEPNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 74


RESULT 9
US-09-284-878-1
; Sequence 1, Application US/09284878
; Patent No. 6342375
; GENERAL INFORMATION:
;  APPLICANT: [illegible], Martha Guerrero
;  APPLICANT: Saldana, Hugo Barrera
;  APPLICANT: [illegible], Jose Maria Viader
;  TITLE OF INVENTION: Genetically Modified Methylotrophic P. pastoris Yeast for the
;  TITLE OF INVENTION: Production and Secretion of the Human Growth Hormone
;  FILE REFERENCE: 1828[illegible]
;  CURRENT APPLICATION NUMBER: US/09/284,878
;  CURRENT FILING DATE: [illegible]
;  PRIOR APPLICATION NUMBER: [illegible]
;  PRIOR FILING DATE: 1997-[illegible]
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 217
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-284-878-1

  Query Match             98.1%;  Score 255;  DB 4;  Length 217;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-29;
  Matches   48;  Conservative    0;  Mismatches    [illegible];  Indels    [illegible];  Gaps    [illegible];

Qy          2 EPEIPLSRLEPNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db         27 EPEIPLSRLEPNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 74


RESULT 10
US-09-424-208-25
; Sequence 25, Application US/09424208
; Patent No. [illegible]
;   GENERAL INFORMATION:
;         APPLICANT: BANYU SYNTHETIC FIBER [illegible]
```

; COMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7647
; INFORMATION FOR SEQ ID NO: 73:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-208-582-73

Query Match 95.8%; Score 265; DB 3; Length 5000;
Best Local Similarity 97.0%; Pred. No. 4.1e-24;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 EEEIPLSKEFDNAFLRRHRQDTAFQSSAYGAALESPGACVPCACQGLDSLVCPSGGGLD 61
EEEIPLSKEFDNAFLRRHRQDTAFQSSAYGAALESPGACVPCACQGLDSLVCPSGGGLD
Db 2 EEEIPLSKEFDNAFLRRHRQDTAFQSSAYGAALESPGACVPCACQGLDSLVCPSGGGLD 61

RESULT 14
; Sequence 14, Application US/09465461
; Patent No. 6514444
; GENERAL INFORMATION:
; APPLICANT: CHAPPEL, Scott
; TITLE OF INVENTION: Human Growth hormone to stimulate hematopoiesis and immune recons
; TITLE OF INVENTION: after hematopoietic stem cell transplantation in humans
; FILE REFERENCE: CHAPPEL 6.2
; CURRENT APPLICATION NUMBER: US/09/465,461
; CURRENT FILING DATE: 1999-12-17
; PRIOR APPLICATION NUMBER: 60/112,668
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Version 3.1
; SEQ ID NO 14
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-465-461-14

Query Match 95.8%; Score 265; DB 4; Length 191;
Best Local Similarity 97.0%; Pred. No. 1.6e-28;
Matches 64; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 EEEIPLSKEFDNAFLRRHRQDTAFQSSAYGAALESPGACVPCACQGLDSLVCPSGGGLD 61
EEEIPLSKEFDNAFLRRHRQDTAFQSSAYGAALESPGACVPCACQGLDSLVCPSGGGLD
Db 48 EEEIPLSKEFDNASLRAHRLRQDTAFDYSPFEAAVLPKEQKYSFLQNPQ 97

RESULT 15
; Sequence 4, Application US/08729560
; Patent No. 5714352
; GENERAL INFORMATION:
; APPLICANT: P.SEN, ET AL.
; TITLE OF INVENTION: Human Growth Hormone
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL, BYERS, DAIN, SIEFFERMAN,
; ADDRESSEE: DEVOE, SEWART & OLSTEIN
; STREET: 4 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/729,560

; FILING DATE: 10-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FERRARO, GREGORY D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 325800.55
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 217 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: UNKNOWN
; MOLECULE TYPE: PROTEIN
US-08-729-560-14

Query Match 95.8%; Score 265; DB 3; Length 217;
Best Local Similarity 97.0%; Pred. No. 1.6e-28;
Matches 64; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 EEEIPLSKEFDNAFLRRHRQDTAFQSSAYGAALESPGACVPCACQGLDSLVCPSGGGLD 61
EEEIPLSKEFDNASLRAHRLRQDTAFDYSPFEAAVLPKEQKYSFLQNPQ 97
Db 121 EEEIPLSKEFDNASLRAHRLRQDTAFDYSPFEAAVLPKEQKYSFLQNPQ 170

RESULT 16
US-08-469-486-51
; Sequence 51, Application US/08/469,486
; Patent No. 5739281
; GENERAL INFORMATION:
; APPLICANT: Thogersen, Hans Christian
; APPLICANT: Reiter, Ilse Lis
; APPLICANT: Lutzkendorf, Michael
; TITLE OF INVENTION: Improved method for the refolding of
; TITLE OF INVENTION: proteins
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version
; SOFTWARE: #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,486
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE: February 4, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul D. Clark
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE:
; TELEFAX: 617 542 8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 217 amino acids
; TYPE: amino acid

Mgi protein - protein search, using sw model

run on: February 5, 2003, 12:42:12 ; Search time 30.17 Seconds
(without alignments)
20.211 Million cell updates/sec

Title: US-09-423-100-1
Perfect score: 260
Sequence: 1 MTTTTTSPTPTNAMLSAdP.....TTTLAYTKKLSYTTLNE 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 9 847 seqs, 1x25x23 residues

Total number of hits satisfying chosen parameters: 9847

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*

1:	[illegible]	260	100.0	19	2	AAY12[illegible]
2:	[illegible]	260	100.0	19	2	AAY12[illegible]
3:	[illegible]	260	100.0	19	2	AAY12[illegible]
4:	[illegible]	260	100.0	19	2	AAY12[illegible]
5:	[illegible]	260	100.0	19	2	AAY12[illegible]
6:	[illegible]	260	100.0	19	2	AAY12[illegible]
7:	[illegible]	260	100.0	19	2	AAY12[illegible]
8:	[illegible]	260	100.0	19	2	AAY12[illegible]
9:	[illegible]	260	100.0	19	2	AAY12[illegible]
10:	[illegible]	260	100.0	19	2	AAY12[illegible]
11:	[illegible]	260	100.0	19	2	AAY12[illegible]
12:	[illegible]	260	100.0	19	2	AAY12[illegible]
13:	[illegible]	260	100.0	19	2	AAY12[illegible]
14:	[illegible]	260	100.0	19	2	AAY12[illegible]
15:	[illegible]	260	100.0	19	2	AAY12[illegible]
16:	[illegible]	260	100.0	19	2	AAW42[illegible]
17:	[illegible]	260	100.0	19	2	AAY12[illegible]
18:	[illegible]	260	100.0	19	2	AAY12[illegible]
19:	[illegible]	260	100.0	19	2	AAW42[illegible]
20:	[illegible]	260	100.0	19	2	AAY12[illegible]
21:	[illegible]	260	100.0	19	2	AAY12[illegible]
22:	[illegible]	260	100.0	19	2	AAY12[illegible]
23:	[illegible]	260	100.0	19	2	AAY12[illegible]
24:	[illegible]	260	100.0	19	2	AA[illegible]
25:	[illegible]	260	100.0	19	2	AA[illegible]
26:	[illegible]	260	100.0	19	2	AAE2[illegible]
27:	[illegible]	260	100.0	19	2	AAY12[illegible]
28:	[illegible]	260	100.0	19	2	AAE2[illegible]
29:	[illegible]	260	100.0	214	7	AAE2[illegible]
30:	[illegible]	260	100.0	219	2	AAE2[illegible]
31:	[illegible]	260	100.0	219	2	AAB[illegible]
32:	[illegible]	260	100.0	219	2	AAW42[illegible]
33:	[illegible]	260	100.0	219	2	AAY12[illegible]
34:	[illegible]	260	100.0	219	2	AAW42[illegible]
35:	[illegible]	260	100.0	219	2	AAY12[illegible]
36:	[illegible]	260	100.0	219	2	AAB[illegible]
37:	[illegible]	260	100.0	219	2	AAW42[illegible]
38:	[illegible]	260	100.0	219	2	AAE2[illegible]
39:	[illegible]	260	100.0	219	2	AA[illegible]
40:	[illegible]	260	100.0	219	2	AAW42[illegible]
41:	[illegible]	260	100.0	219	2	AA[illegible]
42:	[illegible]	260	100.0	219	2	AAE2[illegible]
43:	[illegible]	260	100.0	219	2	AA[illegible]
44:	[illegible]	260	100.0	219	2	AAW42[illegible]
45:	[illegible]	260	100.0	219	2	AAB[illegible]

First, Nos is the number of results predicted by chance to have a score greater than or equal to the score listed, printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	DB ID	Description
1	100.0	19	AAY12[illegible]	Human growth hormo
2	100.0	19	AAY12[illegible]	Human growth hormo
3	100.0	19	AAY128[illegible]	[illegible]protein[illegible]
4	100.0	191	AAW422[illegible]	Human and [illegible] analogue
5	100.0	191	AAE2[illegible]	Human growth hormo
6	100.0	191	AAY12[illegible]	Chimeric protein
7	100.0	192	AAE2[illegible]	Human growth hormo
8	100.0	192	AAW422[illegible]	Human and analogue
9	100.0	262	AAE2[illegible]	Human beta hormone

PN Claim 1; Page 28; 46pp; English.
XX
CC This sequence represents an N-terminal fragment of human growth
CC hormone (hGH) which is a component of a chimeric protein.
CC hGH mini-proinsulin (AAY42657), the hGH portion of the chimeric protein
CC acts as an intramolecular chaperone (IMC) for the insulin precursor,
CC enabling it to fold correctly. A cleavable peptide linker with a
CC C-terminal Arg residue (AAY42657) enables the hGH portion of the
CC chimeric protein to be removed after folding has taken place. Production
CC of recombinant human insulin via an hGH-proinsulin chimeric protein can
CC provide human insulin with correctly linked cysteine bridges with
CC fewer necessary procedural steps, and hence resulting in a higher yield
CC of human insulin. The IMC sequences not only protect insulin sequences
CC from intracellular degradation by a microorganism host, but also promote
CC the folding of the fused insulin precursor, facilitate the solubility of
CC the fusion protein and decrease the intermolecular interactions among
CC the fusion proteins, thus allowing folding of the fused insulin precursor
CC at commercially useful high concentrations. The procedural steps of
CC cyanogen bromide cleavage, oxidative sulphitolysis and related
CC purification steps can thus be eliminated, along with the
CC need for high concentrations of mercaptan or the use of hydrophobic absorbent resins.
XX
SQ Sequence 87 AA;

Query Match 100.0%; Score 2057; DB 4; Length 87;
Best Local Similarity 100.0%; Pred. No. 2e-25;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEPLIPLGPLFLNAMLRAHHLHQLAADTYQEFEEAYIPKEQKYSFLQNP 49
|||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MEPLIPLGPLFLNAMLRAHHLHQLAADTYQEFEEAYIPKEQKYSFLQNP 49

RESULT 2
AAY42657
ID AAY42657 standard; protein; 87 AA.
XX
AC AAY42657;
XX
DT 15-JAN-2000 (first entry)
XX
DE hGH mini proinsulin chimeric protein.
XX
KW Insulin precursor; growth hormone; chaperone; intramolecular;
KW folding; conformation; chimeric protein; cleavable; N-terminal;
KW protein yield.
XX
OS Synthetic.
OS Homo sapiens
XX
RN [1]
XX
RP 1-87
XX
RX MAP: 19q13.2-qter.
XX
RX MAP: 19q13; 98W-98N-99Q-99Y-2.
XX
RA (LINC) LINCOLN GENTECH LTD.
XX
RT Sun C.;
XX
RL WPI; 2000-616869/59.
XX
PN New chimeric proteins containing human growth hormone fragment, used
PN particularly for the production of human insulin
XX
PS Claim 1; Page 28; 46pp; English.
XX
CC This sequence represents an N-terminal fragment of human growth hormone
CC (hGH) which is a component of a chimeric protein (AAY42657) which also
CC has the A human insulin precursor (AAY42656). The hGH portion of the
CC chimeric protein acts as an intramolecular chaperone (IMC) for the
CC insulin precursor, enabling it to fold correctly. A cleavable peptide
CC linker with a C-terminal Arg residue (AAY42657) enables the hGH portion
CC of the chimeric protein to be removed after folding has taken place.
CC Production of recombinant human insulin via an hGH-proinsulin chimeric
CC protein can provide human insulin with correctly linked cysteine bridges
CC with fewer necessary procedural steps, and hence resulting in a higher
CC yield of human insulin. The IMC sequences not only protect insulin
CC sequences from intracellular degradation by a microorganism host, but
CC also promote the folding of the fused insulin precursor, facilitate the
CC solubility of the fusion protein and decrease the intermolecular
CC interactions among the fusion proteins, thus allowing folding of the
CC fused insulin precursor at commercially useful high concentrations. The
CC procedural steps of cyanogen bromide cleavage, oxidative sulphitolysis
CC and related purification steps can thus be eliminated, along with the
CC need for high concentrations of mercaptan or the use of hydrophobic
CC absorbent resins.
XX
SQ Sequence 87 AA;

Query Match 100.0%; Score 2057; DB 4; Length 87;
Best Local Similarity 100.0%; Pred. No. 2e-25;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEPLIPLGPLFLNAMLRAHHLHQLAADTYQEFEEAYIPKEQKYSFLQNP 49
|||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MEPLIPLGPLFLNAMLRAHHLHQLAADTYQEFEEAYIPKEQKYSFLQNP 49

RESULT 3
AAY42660
ID AAY42660 standard; protein; 87 AA.
XX
AC AAY42660;
XX
DT 15-JAN-2000 (first entry)
XX
DE hGH mini proinsulin chimeric protein.
XX
KW Insulin precursor; growth hormone; chaperone; intramolecular;
KW folding; conformation; chimeric protein; cleavable; N-terminal;
KW protein yield.
XX
OS Synthetic.
OS Homo sapiens
XX
RN [1]
XX
RP 1-87
XX
RX MAP: 19q13.2-qter.
XX
RX MAP: 19q13; 98W-98N-99Q-99Y-2.
XX
RA (LINC) LINCOLN GENTECH LTD.
XX
RT Sun C.;
XX
RL WPI; 2000-616869/59.
XX
PN New chimeric proteins containing human growth hormone fragment, used
PN particularly for the production of human insulin
XX
PS Claim 14; Page 30; 46pp; English.
XX
CC This sequence represents a chimeric protein, hGH mini proinsulin.
CC This chimeric protein contains an N-terminal fragment of human growth
CC hormone (hGH) of the sequence given in AAY42657, a cleavable peptide
CC linker (AAY42657), and a human insulin precursor comprising insulin
CC A and B chains (AAY42656). The hGH portion of the chimeric protein acts
CC as an intramolecular chaperone (IMC) for the insulin precursor,
CC enabling it to fold correctly. The C-terminal Arg residue with cleaves the hGH portion of the
CC chimeric protein to be removed after folding has taken place. Production
CC of recombinant human insulin via an hGH-proinsulin chimeric protein can

```
CC   provide human insulin with correctly linked cysteine bridges with
CC   fewer necessary procedural steps, and hence resulting in a higher yield
CC   of human insulin. The IMP sequences not only protect insulin sequences
CC   from intracellular degradation by a microorganism host, but also promote
CC   the folding of the fused insulin precursor, facilitate the solubility of
CC   the fusion protein and decrease the intermolecular interactions amongst
CC   the fusion proteins, thus allowing folding of the fused insulin precursor
CC   at commercially useful high concentrations. The procedural steps of
CC   cyanogen bromide cleavage, oxidative sulphitolysis and related
CC   purification steps can thus be eliminated, along with the use of high
CC   concentrations of mercaptan or the use of hydrophobic absorbent resins.
XX
SQ   Sequence   107 AA;

  Query Match              100.0%;  Score 26[illegible];  DB 2[illegible];  Length 1[illegible];
  Best Local Similarity    100.0%;  Pred. No. 4.8e-2[illegible];
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 1[illegible]
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 1[illegible]


RESULT 4
AAW[illegible]
ID   AAW[illegible] standard; Protein; [illegible] AA.
XX
AC   AAW[illegible];
XX
DT   [illegible] JUN 1998  (first entry)
XX
DE   Human anti-angiogenic peptide 16K hGH Met-[illegible].
XX
KW   Human; anti-angiogenic; prolactin; placental lactogen; hPL; angiogenesis;
KW   growth hormone; hGH; hGH-V; capillary endothelial cell proliferation;
KW   placental vascularisation; pregnancy; treatment; angiogenic disease;
KW   tumour; inhibitor; malignant; angiofibroma; arteriovenous malformation;
KW   arthritis; atherosclerotic plaques; corneal graft neovascularisation;
KW   wound healing; proliferative retinopathy; macular degeneration; trachoma;
KW   granulation; glaucoma; ocular; uveitis; fracture; Osler Weber syndrome;
KW   psoriasis; fibroplasia; scleroderma; Kaposi's sarcoma; vascular adhesion;
KW   tumour; leukaemia; reproductive disorder; contraceptive agent;
KW   gene therapy; pre-eclampsia; intrauterine growth retardation;
KW   placental dysfunction.
XX
OS   Homo sapiens.
XX
PN   WO[illegible]-A1.
XX
PD   [illegible] NOV 1997.
XX
PF   1[illegible] MAY 1997;  97WO-US0[illegible].
XX
PR   [illegible] MAY 1997;  97US-[illegible].
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Martini JA,  Struman I,  Taylor R,  Weiner R;
XX
DR   WPI; 1998-0[illegible]/04.
DR   N-PSDB; AAX[illegible].
XX
PT   New anti-angiogenic peptides - comprise N-terminal fragments of
PT   human placental lactogen, human growth hormone, growth hormone
PT   variant or human prolactin
XX
PS   Claim [illegible]; Page [illegible]; [illegible]pp; English.
XX
CC   This invention describes novel human anti-angiogenic peptides derived
CC   from [illegible] consecutive amino acids selected from the N-terminal end
CC   of human placental lactogen (hPL), human growth hormone (hGH), growth
CC   hormone variant (hGH-V), or human prolactin. Such peptides (i) inhibit
CC   capillary endothelial cell proliferation and [illegible]isation, (ii) inhibit
CC   angiogenesis in [illegible]
CC   least [illegible]
CC   hGH, hPL, prolactin or hGH-V. The [illegible]
CC   than simply a probable alternative [illegible] vascularisation [illegible]
CC   pregnancy. The peptides can be used for treating [illegible]
CC   a subject, for inhibiting tumour [illegible]
CC   modulating vascularisation of a patient's placenta. In particular [illegible]
CC   peptides can be used for preventing or treating [illegible]
CC   angiofibroma, arteriovenous malformations, arthritis [illegible]
CC   arthritis, atherosclerotic plaques, corneal graft neovascularisation,
CC   delayed wound healing, proliferative retinopathy such as diabetic
CC   retinopathy, macular degeneration, granulations such as [illegible]
CC   in haemophilic joints, inappropriate vascularisation in wound [illegible]
CC   such as hypertrophic scars or keloid scars, neovascular glaucoma, [illegible]
CC   tumour, uveitis, non-union fractures, Osler Weber syndrome, [illegible]
CC   pyogenic granuloma, retrolental fibroplasia, scleroderma, [illegible]
CC   Kaposi's sarcoma, trachoma, vascular adhesions, [illegible]
CC   leukaemia, and reproductive disorders such as [illegible]
CC   and choriocarcinoma. They can also be used as contraceptive agents. DNA
CC   encoding the peptides can be used in gene therapy. The measurement of
CC   abnormal levels of N-terminal fragments of hGH, hGH-V [illegible]
CC   can be used in assays for [illegible]
CC   with pre-eclampsia, intrauterine growth retardation and placental
CC   dysfunction.
XX
SQ   Sequence   [illegible]4 AA;

  Query Match              [illegible]
  Best Local Similarity    [illegible]
  Matches   4[illegible];  Conservative    [illegible];  Mismatches    [illegible];  Indels    [illegible];  Gaps    [illegible];

Qy          1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 1[illegible]
              [illegible]
Db          1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 1[illegible]


RESULT 5
AAB[illegible]
ID   AAB[illegible] standard; protein; [illegible] AA.
XX
AC   AAB[illegible];
XX
DT   [illegible] 2001  (first entry)
XX
DE   Human growth hormone [illegible]
XX
KW   Human growth hormone [illegible]
KW   [illegible]
XX
OS   Homo sapiens (human).
XX
PN   EP[illegible]-A.
XX
PD   [illegible]
XX
PF   [illegible]
XX
PR   [illegible]
XX
PA   [illegible]
XX
PI   Ontsuka [illegible]
XX
DR   WPI; [illegible]
XX
PT   New human [illegible] growth hormone [illegible]
PT   [illegible]
PT   dementia, etc.
XX
PS   [illegible]
XX
CC   Human growth hormone [illegible] N-terminal [illegible]
CC   protein [illegible]
```

XX
SQ Sequence 191 AA;

Query Match [illegible] Score 269; DB 14; Length 191;
Best Local Similarity [illegible] Pred. No. 1.4e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49

RESULT 7
AAW92264
ID AAW92264 standard; protein; 192 AA.
XX
AC AAW92264;
XX
DT 06-FEB-1998 (revised 8)
DT 01-NOV-1997 (first entry)
XX
DE Human growth hormone fusion protein for faster secretion
XX
KW Human growth hormone; fusion protein; secretion enhancement;
KW recombinant.
XX
OS Homo sapiens (human).
XX
FN JP10113181-A.
XX
PD 07-JUN-1998.
XX
PF 02-NOV-1996; 96JP-000292.
XX
PR 02-NOV-1996; 96JP-000292.
XX
PA (WAKU) WAKUNAGA PHARM KK.
XX
DR WPI; 1998-339264/29.
DR N-PSDB; AAN90209.
XX
PT Recombinant vector containing fusion protein consisting of human
PT growth hormone derivative linked to leader protein for stabilizing
PT and high yield
XX
PS Disclosure; Pg 11; 20pp; Japanese.
XX
CC The invention consists of a vector which, a fusion protein which is
CC formed by signaling downstream of a promoter, hGH or a part (first
CC terminal by substitution of Met-11 with Leu) and a heterologous protein.
CC Stability of the vector in the host is greatly increased and
CC protein yield is higher.
XX
SQ Sequence 192 AA;

Query Match [illegible] Score 269; DB 14; Length 192;
Best Local Similarity [illegible] Pred. No. 1.4e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49

RESULT 8
AAW92264
ID AAW92264 standard; protein; 192 AA.
XX
AC AAW92264;
XX
DT 06-FEB-1998 (revised 8)
DT 01-NOV-1997 (first entry)
XX
DE Human growth hormone fusion protein for faster secretion
XX
KW Human growth hormone; fusion protein; secretion enhancement;
KW recombinant.
XX
OS Homo sapiens (human).
XX
RN [illegible]
XX
RP 02-NOV-1996; 96JP-000292.
XX
RP 02-NOV-1996; 96JP-000292.
XX
RA (WAKU) WAKUNAGA PHARM KK.
XX
RT Recombinant vector containing fusion protein consisting of human
RT growth hormone derivative linked to leader protein for stabilizing
RT and high yield
XX
RL Disclosure; Pg 11; 20pp; Japanese.
XX
CC The invention consists of a vector which, a fusion protein which is
CC formed by signaling downstream of a promoter, hGH or a part (first
CC terminal by substitution of Met-11 with Leu) and a heterologous protein.
CC Stability of the vector in the host is greatly increased and
CC protein yield is higher.
XX
SQ Sequence 192 AA;

Query Match [illegible] Score 269; DB 14; Length 192;
Best Local Similarity [illegible] Pred. No. 1.4e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49

RESULT 9
AAW92264
ID AAW92264 standard; protein; 192 AA.
XX
AC AAW92264;
XX
DT 06-FEB-1998 (revised 8)
DT 01-NOV-1997 (first entry)
XX
DE Human growth hormone fusion protein for faster secretion
XX
KW Human growth hormone; fusion protein; secretion enhancement;
KW recombinant.
XX
OS Homo sapiens (human).
XX
RN [illegible]
XX
RP 02-NOV-1996; 96JP-000292.
XX
RP 02-NOV-1996; 96JP-000292.
XX
RA (WAKU) WAKUNAGA PHARM KK.
XX
RT Recombinant vector containing fusion protein consisting of human
RT growth hormone derivative linked to leader protein for stabilizing
RT and high yield
XX
RL Disclosure; Pg 11; 20pp; Japanese.
XX
CC The invention consists of a vector which, a fusion protein which is
CC formed by signaling downstream of a promoter, hGH or a part (first
CC terminal by substitution of Met-11 with Leu) and a heterologous protein.
CC Stability of the vector in the host is greatly increased and
CC protein yield is higher.
XX
SQ Sequence 192 AA;

Query Match [illegible] Score 269; DB 14; Length 192;
Best Local Similarity [illegible] Pred. No. 1.4e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49

RESULT 10
AAW92264
ID AAW92264 standard; protein; 192 AA.
XX
AC AAW92264;
XX
DT 06-FEB-1998 (revised 8)
DT 01-NOV-1997 (first entry)
XX
DE Human growth hormone fusion protein for faster secretion
XX
KW Human growth hormone; fusion protein; secretion enhancement;
KW recombinant.
XX
OS Homo sapiens (human).
XX
RN [illegible]
XX
RP 02-NOV-1996; 96JP-000292.
XX
RP 02-NOV-1996; 96JP-000292.
XX
RA (WAKU) WAKUNAGA PHARM KK.
XX
RT Recombinant vector containing fusion protein consisting of human
RT growth hormone derivative linked to leader protein for stabilizing
RT and high yield
XX
RL Disclosure; Pg 11; 20pp; Japanese.
XX
CC The invention consists of a vector which, a fusion protein which is
CC formed by signaling downstream of a promoter, hGH or a part (first
CC terminal by substitution of Met-11 with Leu) and a heterologous protein.
CC Stability of the vector in the host is greatly increased and
CC protein yield is higher.
XX
SQ Sequence 192 AA;

Query Match [illegible] Score 269; DB 14; Length 192;
Best Local Similarity [illegible] Pred. No. 1.4e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49

XX
KW Human anti-angiogenic protein fragment; HAP; N-terminal; anti-angiogenic;
KW N-terminal domain; human placental vascularisation; pregnancy; treatment; angiogenic disease;
KW tumour; solid tumour; malignancy; anti-fibrotic; arteriovenous malformation;
KW arthritis; atherosclerotic plaque; corneal graft neovascularisation;
KW wound healing; proliferative retinopathy; macular degeneration; trachoma;
KW granulation; atherosclerosis; ocular uveitis; fracture; Osler Weber syndrome;
KW psoriasis; HIV promoter; sickle-cell; Kaposi's sarcoma; vascular adhesion;
KW uterine leukaemia; reproductive disorder; contraceptive agent;
KW gene therapy; pre-eclampsia; intrauterine growth retardation;
KW placental dysfunction.
XX
OS Homo sapiens.
XX
PN WO200240541-A1.
XX
PD 23-MAY-2002.
XX
PF 16-NOV-2001; 2001WO-EP013417.
XX
PR 16-NOV-2000; 2000US-0713999P.
XX
PA (UYPT-) UNIV ALBI ENTA.
XX
PI Martin A, Graham CH, Lye SJ, Kniss DA, Walker JJ;
XX
DR WPI; 2002-521242/56.
DR N-PSDB; AAX17744.
XX
PT New anti-angiogenic peptides comprise N-terminal fragments of
PT human placental vascular growth factor with activity to
PT inhibit angiogenesis in human placenta.
XX
PS Example 1; Fig 1; 47pp; English.
XX
CC The invention describes novel human anti-angiogenic peptides derived
CC from 10 to 100 consecutive amino acids selected from the N-terminal end
CC of human placental growth factor (hPGF), human or mouse vascular endothelial
CC growth factor (hVEGF or mVEGF), or human placenta growth factor (hPlGF).
CC The peptides (I) inhibit capillary endothelial cell proliferation and organisation (ii) inhibit
CC angiogenesis in chick chorioallantoic membrane and (iii) binds to at
CC least one specific receptor which does not bind an intact full length
CC hVEGF, mVEGF, placenta growth factor or hVEGF-D. The invention also describes a method for
CC diagnosing a placental abnormality of placental vascularisation during
CC pregnancy. The peptides can be used for treating an angiogenic disease in
CC a subject, for inhibiting tumour formation or growth in a patient or for
CC modulating vascularisation of a patient's placenta. In particular, the
CC peptides can be used for preventing or treating solid malignant tumours,
CC anti-fibroma, arteriovenous malformation, arthritis such as rheumatoid
CC arthritis, atherosclerotic plaques, corneal graft neovascularisation,
CC haemophilic joints, inappropriate vascularisation in wound healing
CC such as hyperplastic scars or keloid scars, fracture repair, ocular
CC uveitis, ocular inflammatory diseases, sickle-cell anaemia, psoriasis,
CC psoriatic plaque, retrolental fibroplasia, scleroderma, solid tumours,
CC Kaposi's sarcoma, trachoma, vascular adhesions, chronic inflammatory diseases,
CC arthritis, and reproductive disorders such as follicular and luteal cysts,
CC and hydatidiform moles. They can also be used as contraceptive agents. DNA
CC encoding the peptides can be used in gene therapy. The polypeptides or
CC DNA encoding the peptides can be used in pharmaceutical compositions
CC which can be used in assays to determine if cancer, developmental abnormality
CC such as pre-eclampsia, intrauterine growth retardation, and placental
CC dysfunction, etc.
XX
SQ Sequence 152 AA;

Query Match [illegible]
Best Local Similarity [illegible]
Matches [illegible]

RESULT 1
AAP[illegible]
ID AAP[illegible] standard; protein; [illegible] AA.
XX
AC AAP[illegible];
XX
DT [illegible] (first entry).
XX
DE Human beta-2 with [illegible]
KW Human beta-2 with [illegible]
KW Gestation hormone; [illegible]
XX
OS Homo sapiens (human).
XX
FH Key Location/Qualifiers
FT Peptide [illegible]
FT Peptide [illegible]
FT Peptide [illegible]
XX
PN [illegible] A.
XX
PD [illegible] A1.
XX
PF [illegible]
XX
PR [illegible]
XX
DR WPI; [illegible]
XX
PT New [illegible]
PT [illegible]
PT [illegible]
XX
PS [illegible]
XX
CC Fusion protein [illegible] human beta-2 with [illegible]
CC N-terminal end (1st [illegible]), [illegible] and [illegible]
CC thrombin recognition sites and human beta-2 with [illegible]
CC at the C-terminus. Beta-2 NGF can be used [illegible]
CC and other beta-2 disorders and can be produced [illegible]
CC protein by incubation with thrombin (see AAN[illegible] and AA[illegible]).
XX
SQ Sequence [illegible] AA;

Query Match [illegible]
Best Local Similarity [illegible]
Matches [illegible]

RESULT 2
AAP[illegible]
ID AAP[illegible] standard; protein; [illegible] AA.
XX
AC AAP[illegible];
XX
DT [illegible] (first entry).
XX
DE Human beta-2 with [illegible]
XX

```
RW   ??? NTF; ?,???,??? as.
XX
 S   ??? sequences.
XX
FH   Key             Location/qualifiers
FT   ??? in          24?..2??
XX
FN   JP6?,?48?-A.
XX
 ?   ?? Shiritsu??.
XX
 ?   ??MAB-1??;   ??JP-???4??.
XX
PF   ??? MAB 1??5;   ??JP-???4??.
XX
PA   (?ISO?) ?.?HKA ?
XX
PR   WPI; ???-?8?6?/44.
XX
PT   Gene segment of human nerve growth factor 1 ?? used in produ?? of
PT   NGF producing recombinant Escherichia strain.
XX
PS   Claim ?2; Page 48?; ?1pp; Japanese.
XX
CC   The protein is a direct translation of the upstream tryptophan
CC   promoter, operator lacking its attenuation sequence and human
CC   beta NGF sequence. The product may be efficiently expressed from a
CC   transformed E.coli expression system.
CC   See also: AAN??16?.
XX
SQ   Sequence   2?2 AA;

  Query Match             100.0%;  Score 2??;  DB 7;  Length 2??;
  Best Local Similarity   100.0%;  Pred. No. 1.?e-24;
  Matches  ??;  Conservative   ??;  Mismatches   ??;  Indels   ??;  Gaps   ??;

Qy       1 MSFIFLTAILFLNNAISSKADSQSVSAGPGNSLRQKLAFDTYQEIEEAYVHIEKLKSNLKNF 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MSFIFLTAILFLNNAISSKADSQSVSAGPGNSLRQKLAFDTYQEIEEAYVHIEKLKSNLKNF 60


RESULT ??
AAN??16?
ID   AAN??16? standard; protein; ??? AA.
XX
AC   AAN??16?;
XX
DT   ??-???-????  (first entry)
XX
DE   Human ?-NGF (first entry).
XX
KW   ?-NGF; nerve growth factor; beta subunit-precursor; protein
KW   fusion; ??? gene; ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ???.
XX
OS   Homo sapiens.
XX
RN   [1]
RP   1-???
RX   MEDLINE; ????????.
XX
RA   ??? ?.?.;
XX
RT   "Production of human beta-nerve growth factor";
XX
RL   ??? ???? ????  (first entry)
XX
DR   WPI; ????-?????/??.
XX
PT   Partial cDNA sequence of human beta nerve growth factor;
PT   ??? amino acid sequence.
XX
PS   Claim ?; Page ??; ?pp; Japanese.
XX
CC   This sequence represents the human beta nerve growth factor.
XX
SQ   Sequence   ??? AA;

  Query Match             ???.?%;  Score ???;  DB ?;  Length ???;
  Best Local Similarity   ???.?%;  Pred. No. ?.?e-??;
  Matches  ???;  Conservative   ??;  Mismatches   ??;  Indels   ??;  Gaps   ??;

Qy       1 MSFIFLTAILFLNNAISSKADSQSVSAGPGNSLRQKLAFDTYQEIEEAYVHIEKLKSNLKNF 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MSFIFLTAILFLNNAISSKADSQSVSAGPGNSLRQKLAFDTYQEIEEAYVHIEKLKSNLKNF 60


RESULT ??
AAN??16?
ID   AAN??16? standard; protein; ??? AA.
XX
AC   AAN??16?;
XX
DT   ??-???-????  (first entry)
XX
DE   Fusion protein of ?-nerve stimulatory factor 2 and B-cell
DE   differentiation factor.
XX
KW   ?-nerve stimulatory factor 2; interleukin-?; ?-cell differentiation
KW   interleukin-?; ??? gene; ??? protein.
XX
OS   Homo sapiens.
XX
RN   [1]
RP   1-???
RX   MEDLINE; ????????.
XX
PA   ???-JAN-????.
XX
PT   ?? ??? ????;   ??JP-???-???.
XX
PR   ?? ??? ????;   ??JP-???-???.
XX
PA   (?-Y-?) ?.?K-?-??.
XX
DR   WPI; ????-???????/??.
DR   N-PSDB; AA?.???.
XX
PT   Preparation of human B-cell differentiation factor - transcription
PT   sequence-encoding DNA fragment, and its production using
PT   specified amino acid sequence.
XX
PS   Claim ??; Page ??; ??pp; Japanese.
XX
CC   The protein is produced by E.coli DNA expression. SCI (??, ???-???)
CC   considers DSM ?? (??) and ??? ??? ??? ??? ??? ??? ??? ??? ??? ???
CC   See also AAR??411 and AAR??413.
XX
SQ   Sequence   ??? AA;

  Query Match             ???.?%;  Score ???;  DB ?;  Length ???;
  Best Local Similarity   ???.?%;  Pred. No. ?.?e-??;
  Matches  ???;  Conservative   ??;  Mismatches   ??;  Indels   ??;  Gaps   ??;

Qy       1 MSFIFLTAILFLNNAISSKADSQSVSAGPGNSLRQKLAFDTYQEIEEAYVHIEKLKSNLKNF 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MSFIFLTAILFLNNAISSKADSQSVSAGPGNSLRQKLAFDTYQEIEEAYVHIEKLKSNLKNF 60


RESULT ??
AAR??413
ID   AAR??413 standard; protein; 141 AA
XX
AC   AAR??413;
```

XX
PI Pr [illegible] (first entry)
XX
DE Subunit of B cell stimulatory factor 2 (BSF-2).
XX
KW B cell stimulatory factor-2; interleukin 6.
XX
OS Homo sapiens.
XX
RN [1]
RA [illegible] A.
XX
RL [illegible]
XX
RF [illegible]
XX
RF [illegible]
XX
RA [illegible]
XX
RT Will [illegible]
RL Nature 324:73-76.
XX
CC [illegible] cDNA clone encoding human BSF-2 was specified DNA
CC sequence assembly, by recombinant DNA technique, also provided
CC specified amino acid sequence.
XX
CC Recommended name is IFNbeta-2 precursor.
XX
CC Human protein exhibits this protein can be used with DNA encoding B-cell
CC differentiation factor (BCDF, and ligated in an expression vector to
CC produce a fusion protein.
CC See also AAB59411.
XX
SQ Sequence 184 AA;

Query Match [illegible] Score [illegible]; DB [illegible] Length 184;
Best Local Similarity [illegible]; Pred. No. [illegible];
Matches 44; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy [illegible]
Db [illegible]

RESULT 2
AAB59411
ID AAB59411 standard; protein; 184 AA.
XX
AC AAB59411;
XX
DT 28-NOV-1990 (first entry)
XX
DE Human growth hormone with cholecystokinin site.
XX
KW Human growth hormone; cholecystokinin site.
XX
OS Homo sapiens.
XX
RN [1]
XX
RP 1-184
XX
RX DOI [illegible]
XX
RA [illegible]
XX
RA [illegible]
XX
RA [illegible]
XX
RT Will [illegible]
RL Nature 324:73-76.
XX
CC Polypeptide with a structure [illegible]
CC [illegible] growth hormone [illegible]
CC sequence with specified DNA sequence.
XX
FH Key Location/Qualifiers
XX
CC The polypeptide (AAB59411) with a sequence [illegible]
CC (AAN81005) encoding it are claimed. A MAb specific [illegible]
CC The N terminal [illegible] polypeptide also [illegible]
CC [illegible] and [illegible] with the polypeptide [illegible]
CC production kits and is useful [illegible]
CC and defined for use as a kit [illegible]
XX
SQ Sequence [illegible] AA;

Query Match [illegible] Score [illegible]; DB [illegible] Length [illegible];
Best Local Similarity [illegible]; Pred. No. [illegible];
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy [illegible]
Db [illegible]

Search completed: [illegible]
Job time : [illegible] secs

Run on: February 5, 2003, 12:42:21 ; Search time 5.2744 Seconds
(without alignments)
[illegible] Million cell updates/sec

Title: US-09-423-100-1
Perfect score: 260
Sequence: 1 MEFTTTLSPTLNAMIPAAS.....IIIIAYIKE_SYSTL_NF 44

Scoring table: BL_SUM62
Gapop 10.0 , Gapext 0.5

Searched: [illegible] seqs, [illegible] residues

[illegible] scores with SUMB parameters: 26.23

Minimum Seq Length:
Maximum Seq Length: [illegible]

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

The N, is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB	ID	Description
1	260	100.0	217	1	S10810 somatotropin l pro
2	237	91.2	217	2	S19731 somatotropin - pig
3	228	87.7	217	1	S10810V somatotropin 2 pro
4	228	87.7	217	2	S19732 somatotropin 2 pro
5	223	85.8	217	2	I55267 Chorionic somatomam
6	220	84.6	217	2	S19731 somatotropin - pig
7	201	77.3	217	2	I67409 Chorionic somatomam
8	197	75.8	217	2	A27449 Growth hormone-pit
9	197	75.8	217	2	I67850 Growth hormone-pit
10	197	75.8	217	2	B31149 Growth hormone-pit
11	170.5	65.6	210	1	S19851 somatotropin precu
12	169.5	65.2	190	2	PN 1460 somatotropin - sei
13	169.5	65.2	191	2	S19852 somatotropin - Ani
14	169.5	65.2	191	2	A27429 somatotropin - Afr
15	169.5	65.2	191	2	JS0429 somatotropin - Afr
16	169.5	65.2	220	2	S25811 somatotropin precu
17	169.5	65.2	210	1	S19851 somatotropin precu
18	169.5	65.2	220	2	A47147 somatotropin precu
19	169.5	65.2	220	2	S49463 somatotropin precu
20	169.5	65.2	220	2	A40059 somatotropin - sei
21	169.5	65.2	220	2	JH0542 somatotropin precu
22	169.5	65.2	216	2	A27744 somatotropin precu
23	169.5	65.2	191	2	A25494 somatotropin - sei
24	156	60.0	216	2	JH0514 somatotropin precu
25	154	59.2	191	2	A10426 somatotropin - sei
26	149	57.3	191	2	JN0467 somatotropin - sei
27	144	55.4	190	2	S21753 somatotropin - Rhe
28	144	55.4	210	2	A25494 somatotropin precu
29	144	55.4	220	2	A40059 somatotropin - sei
30	144	55.4	217	1	S10 [illegible]
31	143	55.0	217	2	S19 [illegible]
32	141	54.2	217	1	S10 [illegible]
33	141	54.2	217	2	S19 [illegible]
34	138	53.1	217	2	S19 [illegible]
35	138	53.1	217	2	S19 [illegible]
36	137	52.7	217	2	A27 [illegible]
37	137	52.7	217	2	S19 [illegible]
38	137	52.7	217	2	S19 [illegible]
39	137	52.7	217	2	S19 [illegible]
40	136	52.3	190	2	A40 [illegible]
41	135	51.9	217	2	S19 [illegible]
42	135	51.9	217	2	I [illegible]
43	135	51.9	217	2	S19 [illegible]
44	135	51.9	217	2	S19 [illegible]
45	135	51.9	217	2	S19 [illegible]

ALIGNMENTS

RESULT 1
S10810
SOMATOTROPIN I PRECURSOR - PIG
N;Alternate names: growth hormone I precursor; pituitary growth hormone I precursor
N;Contains: growth hormone I (1-191); growth hormone I (1-217)
C;Species: Sus scrofa (pig)
C;Date: 21-Apr-1991 #sequence_revision 21-Apr-1991 #text_change [illegible]
C;Accession: A10731; A10732; A10733; A10734; A10735; A10736; A10737; A10738
R;DeNoto, F.M., Moore, D.D., Goodman, H.M.
Nucleic Acids Res. 9, 3719-3730 (1981)
A;Title: Human growth hormone DNA sequence and its expression in cultured kidney cells
A;Reference number: A10731; MUID:82001; PMID:7031
A;Accession: A10731
A;Molecule type: DNA
A;Residues: 1-217 <DEN
A;Cross-references: GB:V; [illegible]
A;Note: the 20K short [illegible] residues [illegible]
R;Boime, I.Y.; Ilan, J.; Shafer, J.A.; Hsueh, A.J.W.
Genomics 4, 479-497 (1989)
A;Title: The human growth hormone gene family: structure of the sequences
A;Reference number: A31455; MUID:89[illegible]; PMID:[illegible]
A;Accession: A31455
A;Molecule type: DNA
A;Residues: 1-217 <BOI
A;Cross-references: [illegible]
R;Roskam, W.; Reuvens, [illegible]
Nucleic Acids Res. [illegible]
A;Title: Molecular cloning and characterization of the human growth hormone gene
A;Reference number: A10733; MUID:[illegible]; PMID:[illegible]
A;Accession: A10733
A;Molecule type: mRNA
A;Residues: 1-217 <ROS
A;Cross-references: [illegible]
A;Note: 20 K [illegible]
R;Martial, J.A.; Baxter, J.D.; Goodman, H.M.
Science 205, [illegible]
A;Title: Human growth hormone: complete nucleotide sequence and [illegible]
A;Reference number: A01143; MUID:[illegible]; PMID:[illegible]
A;Accession: A01143
A;Molecule type: mRNA
A;Residues: 1-217 <MAR
R;[illegible]
Arch. Biochem. Biophys. [illegible]
A;Title: Human pituitary growth hormone. XXX. The primary [illegible]
A;Reference number: A[illegible]; MUID:[illegible]; PMID:[illegible]
A;Accession: A00[illegible]
A;Molecule type: protein

A;Residues: 1-[illegible]
E;Nikkhil, R.G.
Nature New Biol., 244, 114-116, 1973
A;Title: Revised primary structure for human growth hormone.
A;Reference number: A93397; MUID:73211926; PMID:4579046
A;Accession: A93397
A;Molecule type: protein
A;Residues: 27-191<NIA>
E;Nikkhil, R.G.; Bosnia, M.J.; Sarma, P.S.; Pesonthim, U.Y.; Greene-Al, L.J.
Proc. Natl. Acad. Sci. U.S.A., 68, 866-869, 1971
A;Title: Sequence of pituitary and placental lactogenic and growth hormones: evolution
A;Reference number: A93730; MUID:71139769; PMID:5278598
A;Accession: A93730
A;Molecule type: protein
A;Residues: 1[illegible]-120;157-159<NIA>
E;Nikkhil, R.G.
Endocrinology and Carcinogenesis, Proc. Fourth Tenovus Workshop Probertin, Griffiths, K.,
A;Title: The chemistry of the human lactogenic hormones.
A;Reference number: A94427
A;Contents: annotation; somatotropin revision
A;Authors: Li, C.H.; Liu, T.S.; Li, C.H.
Int. J. Pept. Protein Res. 4, 281-287, 1972
A;Title: Sequence comparison of human pituitary growth hormones, human chorionic somatom
A;Reference number: A91764; MUID:73092029; PMID:4679454
A;Accession: A91764
A;Molecule type: protein
A;Residues: 27-217<HEW>
[illegible] Y.; Li, C.H.; Papkoff, H.; Liu, T.S.
Biochem. Biophys. Res. Commun., 42, 511-516, 1971
A;Title: The 20,000 molecular weight variant of human growth hormone: I. Purification of the 20K
A;Reference number: A92217; MUID:72017076; PMID:4107475
A;Contents: somatotropin, 20K short variant
A;Accession: A92217
A;Molecule type: protein
A;Residues: 1-[illegible]<HEW>
E;Chapman, G.E.; Roberts, K.M.; Brittain, T.; Bradshaw, R.A.; Bates, [illegible] Turner, J.F.
J. Biol. Chem., 256, 2395-2401, 1981
A;Title: The 20,000 molecular weight variant of human growth hormone. Preparation and
A;Reference number: A92311; MUID:81117961; PMID:7462247
A;Contents: somatotropin, 20K short variant
A;Accession: A92311
A;Molecule type: protein
A;Residues: 27-67;70-191<HUA>
E;Stark, R.N.P.; Searcy, R.F.; Lewis, L.J., I [illegible], D.J.
J. Protein Chem., 2, 425-434, 1983
A;Title: Human growth hormone peptide 1-43: isolation from pituitary glands.
A;Reference number: A61466
A;Accession: A61466
A;Molecule type: protein
A;Residues: 27-67<SIN>
A;Note: growth hormone 5K peptide has insulin potentiating activity; its physiological p
E;Tsiotsos, V.M.; De Bus, L.; N[illegible]
Biol. Chem. Hoppe Seyler, 371, 423-431, 1990
A;Title: Identification of the aspartimide structure in a previously reported peptide
A;Reference number: S00084; MUID:90384745; PMID:2378679
A;Accession: S00084
A;Molecule type: protein
A;Residues: 27-44,47,49-67,70-191<PUB>
E;de Vos, A.M.; Ultsch, M.; Kossiakoff, A.A.
Science, 255, 306-312, 1992
A;Title: Human growth hormone and extracellular domain of its receptor: crystal structu
A;Reference number: A41428; MUID:92196022; PMID:1549776
A;Contents: annotation; X-ray crystallography, 2.9 angstroms
A;Note: the structure of the complex with growth hormone receptor is described
E;Chang, C.N.; Rey, M.; Bochner, B.; Heyneker, H.; Gray, G.
Gene, 55, 189-196, 1987
A;Title: Periplasmic production of correctly processed human growth hormone in Escheri
A;Reference number: [illegible]; MUID:88161709; PMID:3122261
A;Accession: [illegible]
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-[illegible]
A;Cross-reference: GB:M13438; NID:g181[illegible]; PIDN:AAA51[illegible]; PID:g181[illegible]

C;Comment: The gene for this hormone is transcribed only in somat[illegible]
C;Comment: About [illegible] is the 22K [illegible]
C;Genetics:
A;Gene: GDB:GH1
A;Cross-reference: GDB:[illegible]; OMIM:[illegible]
A;Map position: 17p[illegible]
A;Introns: 1/1; 2/2; 3/3; 4/4
C;Superfamily: prolactin
C;Keywords: alternative splicing; hormone; pituitary
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-217/Product: somatotropin 1, [illegible] #status experimental <MAT>
F;27-67,70-191/Product: growth hormone 5K peptide #status experimental <5KP>
F;27-57,73-217/Product: somatotropin 1, short form #status experimental <S-S>
F;79-191,211-219/Disulfide bonds: #status experimental

Query Match [illegible]; Score [illegible]; DB [illegible]; Length [illegible]
Best Local Similarity [illegible]; Pred. No. [illegible]
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible]

Qy [illegible] FFIPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQ 60
[illegible]
Db [illegible] FFIPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQ 74

RESULT 2
167410
somatotropin - rhesus monkey
N;Alternate names: growth hormone
C;Species: Macaca mulatta (rhesus monkey)
C;Date: 31-May-1990 #sequence_revision 31-May-1990 #text_change 16-Jul-1993
C;Accession: I67410; A05044
R;Golos, T.G.; Hattori, M.; Fisher, J.M.; Bowers, P.D.
Endocrinology 133, 1744-1752, 1993
A;Title: Cloning of four growth hormone/chorionic somatomammotropin genes
A;Reference number: I54267; MUID:94098424; PMID:8104617
A;Accession: I67410
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-217<RES>
A;Cross-reference: GB:L19556; NID:g[illegible]; PIDN:AAA[illegible]; PID:g[illegible]
E;Li, C.H.; Chung, D.; Lahm, H.W.; Stein, S.
Arch. Biochem. Biophys. 215, 287-291, 1982
A;Title: The primary structure of monkey pituitary growth hormone
A;Reference number: A27294; MUID:83012747; PMID:6809859
A;Accession: A27294
A;Molecule type: protein
A;Residues: 27-[illegible]
A;Note: the monkey species is not identified in the reference
E;Raben, M.S.
Science 125, 883-884, 1957
A;Title: Preparation of growth hormone in crystalline form and its [illegible]
A;Reference number: A44771
A;Contents: annotation; identification of 2 more isoforms
C;Superfamily: prolactin

Query Match [illegible]; Score [illegible]; DB [illegible]; Length [illegible]
Best Local Similarity [illegible]; Pred. No. [illegible]
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible]

Qy [illegible] FFIPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQ 60
[illegible]
Db [illegible] FFIPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQ 71

RESULT 3
S18DUV
somatotropin 2 precursor - human
N;Alternate names: growth hormone 2; growth hormone variant; hGH-V; hormone, somatot
N;Contains: somatotropin 2; [illegible] somatotropin
C;Species: Homo sapiens (man)
C;Date: 17-Dec-1986 #sequence_revision 16-Jul-1993 #text_change 31-Jul-2
C;Accession: S18DUV; P01242; A02711
R;Chen, E.Y.; Liao, Y.C.; Smith, D.H.; Barrera-Saldana, H.A.; Gelinas, R.E.; Seeburg

RESULT 7
A26419
N;Alternate names: growth hormone
C;Species: Macaca mulatta (rhesus macaque)
C;Date: 31-May-1990 #sequence_revision 31-May-1990 #text_change 16-Jul-1999
C;Accession: A26419
R;Bolos, J.G.; Durning, M.; Fisher, J.M.; Fowler, P.D.
Endocrinology 126, 1744-1752, 1990
A;Title: Cloning of four growth hormone/chorionic somatomammotropin-related complementary
A;Reference number: 154267; MUID:90308724; PMID:8404617
A;Accession: I67411
A;Molecule type: mRNA
A;Residues: 1-217 <BOL>
A;Cross-references: GB:M57757; NID:g339441; PIDN:AAA36641.1; PID:g339442
A;Experimental source: placenta
R;Bolos, J.G.; Durning, M.; Fisher, J.M.; Fowler, P.D.
Endocrinology 126, 1744-1752, 1990
A;Title: Cloning of four growth hormone/chorionic somatomammotropin-related complementary
A;Reference number: 154267; MUID:90308724; PMID:8404617
A;Accession: I67410
A;Molecule type: mRNA
A;Residues: 1-217 <BOL>
A;Cross-references: GB:M57758; NID:g339443; PIDN:AAA36642.1; PID:g339444
A;Experimental source: placenta
R;Bolos, J.G.; Durning, M.; Fisher, J.M.; Fowler, P.D.
Endocrinology 126, 1744-1752, 1990
A;Title: Cloning of four growth hormone/chorionic somatomammotropin-related complementary
A;Reference number: 154267; MUID:90308724; PMID:8404617
A;Accession: I67411
A;Molecule type: mRNA
A;Residues: 1-217 <BOL>
A;Cross-references: GB:M57757; NID:g339441; PIDN:AAA36641.1; PID:g339442
A;Experimental source: placenta
R;Bolos, J.G.; Durning, M.; Fisher, J.M.; Fowler, P.D.
Endocrinology 126, 1744-1752, 1990
A;Title: Cloning of four growth hormone/chorionic somatomammotropin-related complementary
A;Reference number: 154267; MUID:90308724; PMID:8404617
A;Accession: I67410
A;Molecule type: mRNA
A;Residues: 1-217 <BOL>
A;Cross-references: GB:M57758; NID:g339443; PIDN:AAA36642.1; PID:g339444
A;Experimental source: placenta
C;Species: Macaca mulatta (rhesus macaque)
C;Date: 31-May-1990 #sequence_revision 31-May-1990 #text_change 16-Jul-1999
C;Accession: I67409
R;Bolos, J.G.; Durning, M.; Fisher, J.M.; Fowler, P.D.
Endocrinology 126, 1744-1752, 1990
A;Title: Cloning of four growth hormone/chorionic somatomammotropin-related complementary
A;Reference number: 154267; MUID:90308724; PMID:8404617
A;Accession: I67409
A;Molecule type: mRNA
A;Residues: 1-217 <BOL>
A;Cross-references: GB:M57759; NID:g339445; PIDN:AAA36643.1; PID:g339446
C;Superfamily: prolactin

Query Match 77.6%; Score 201; DB 2; Length 217;
Best Local Similarity 74.5%; Pred. No. 1.6e-17;
Matches 143; Conservative 8; Mismatches 41; Indels 10; Gaps 3;

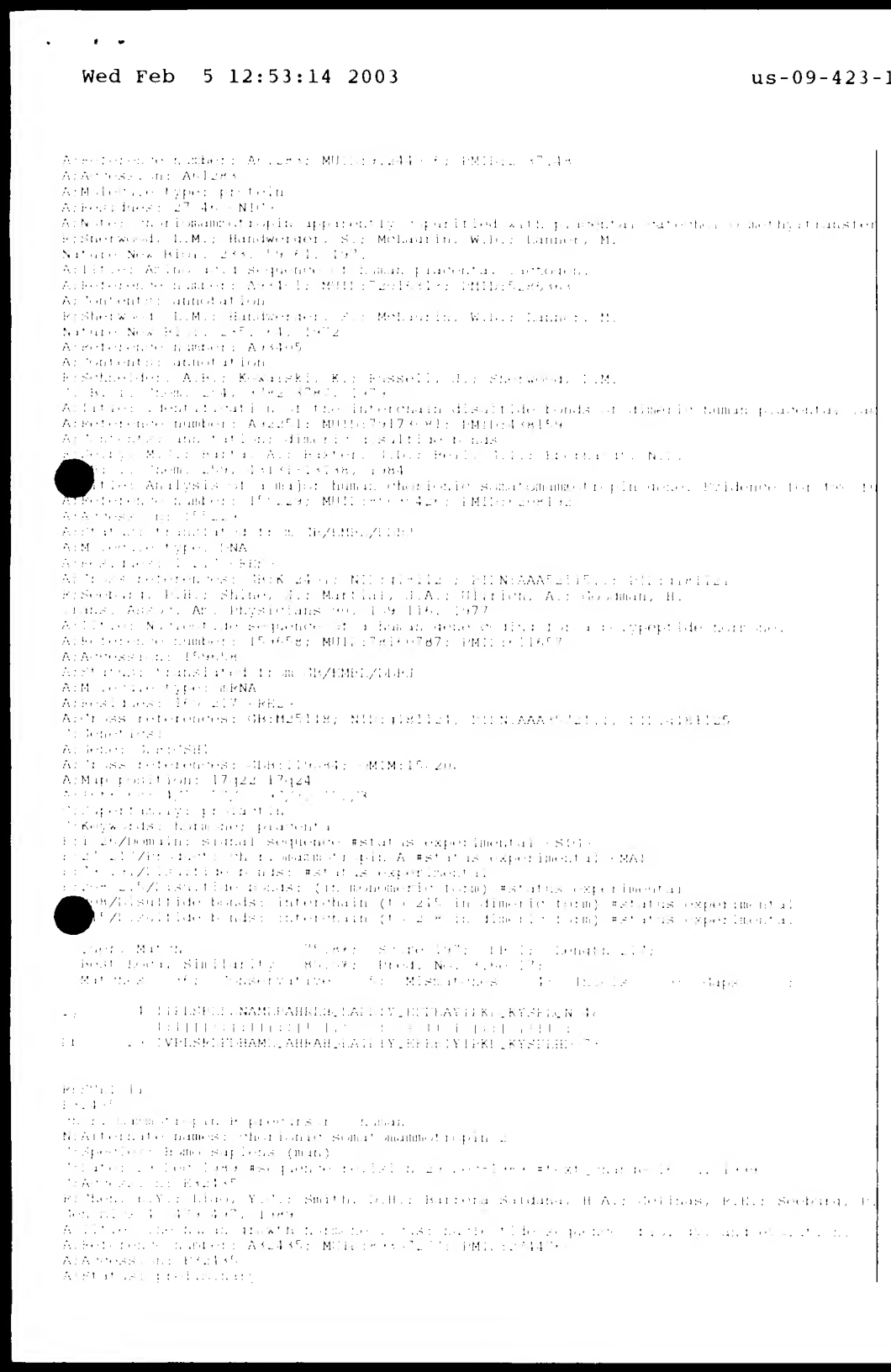
Qy 1 MATGSRTSLLLAFGLLCLPWLQEGSAFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPTPSNREETQQKSNLELLRISLLLIQSWLEPVQFLRSVFANSLVYGASDSNVYDLLKDLEEGIQTLMGRLEDGSPRTGQIFKQTYSKFDTNSHNDDALLKNYGLLYCFRKDMDKVETFLRIVQCRSVEGSCGF 201
Db 1 MATGSRTSLLLAFGLLCLPWLQEGSAFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPTPSNREETQQKSNLELLRISLLLIQSWLEPVQFLRSVFANSLVYGASDSNVYDLLKDLEEGIQTLMGRLEDGSPRTGQIFKQTYSKFDTNSHNDDALLKNYGLLYCFRKDMDKVETFLRIVQCRSVEGSCGF 201

RESULT 8
A26419
C;Species: Homo sapiens (man)
C;Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 16-Jul-1999
C;Accession: A26419
R;Hall, L.R.; Kirkham, J.E.; Burnstone, M.S.; Chen, E.Y.; Seeburg, P.H.; Eberhardt, N.L.; Ba
DNA 6, 59-70, 1987
A;Title: The human growth hormone gene locus: structure, expression, and allelic variati
A;Reference number: A26419; MUID:87191285; PMID:3555880
A;Accession: A26419
A;Molecule type: DNA
A;Residues: 1-217 <HAL>
C;Superfamily: prolactin
F;1,26/Gene: human somatotropin precursor #status predicted <SIG>
F;1-217/Gene: human somatotropin; hGH variable #status predicted <MAT>

Query Match 75.8%; Score 197; DB 2; Length 217;
Best Local Similarity 74.5%; Pred. No. 1.2e-16;
Matches 143; Conservative 8; Mismatches 41; Indels 10; Gaps 3;

Qy 1 MATGSRTSLLLAFGLLCLPWLQEGSAFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPTPSNREETQQKSNLELLRISLLLIQSWLEPVQFLRSVFANSLVYGASDSNVYDLLKDLEEGIQTLMGRLEDGSPRTGQIFKQTYSKFDTNSHNDDALLKNYGLLYCFRKDMDKVETFLRIVQCRSVEGSCGF 201
Db 1 MATGSRTSLLLAFGLLCLPWLQEGSAFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPTPSNREETQQKSNLELLRISLLLIQSWLEPVQFLRSVFANSLVYGASDSNVYDLLKDLEEGIQTLMGRLEDGSPRTGQIFKQTYSKFDTNSHNDDALLKNYGLLYCFRKDMDKVETFLRIVQCRSVEGSCGF 201

RESULT 9
LCHUG
Chorionic somatomammotropin A precursor (placental lactogen)
N;Alternate names: chorionic somatomammotropin pit; placental lactogen
C;Species: Homo sapiens (man)
C;Date: 25-Jul-1986 #sequence_revision 25-Jul-1986 #text_change 16-Jul-1999
C;Accession: A24455; A24456; A24457; A24458; A24459; A24460; A24461; A24462
R;Chen, E.Y.; Liao, Y.C.; Smith, D.H.; Barrera-Saldana, H.A.; Gelinas, R.E.; Seeburg,
Genomics 4, 479-497, 1989
A;Title: The human growth hormone gene family: nonexpressed sequences within a 66-kb
A;Reference number: A24455; MUID:89281744; PMID:2744260
A;Accession: A24456
A;Molecule type: DNA
A;Residues: 1-217 <CHE>
A;Cross-references: GB:M13438; NID:g339447; PIDN:AAA36644.1; PID:g339448
R;Goodman, H.M.; deNoto, F.; Fiddes, J.C.; Barkwell, P.A.; Hall, L.; Moore, R.; Ullrich,
In Molecular and Cellular Endocrinology of the Fetus and Newborn, pp. 3-20, A.R. Liss,
A;Reference number: A94422
A;Accession: A94422
A;Molecule type: mRNA
A;Residues: 1-217 <GOO>
R;Tanaka, M.; Maeda, N.; Watanabe, M.; Yamamoto, M.; Shimizu, K.; Kato, T.; Nagata,
Biochem. Int. 7, 287-292, 1983
A;Title: cDNA cloning of human chorionic somatomammotropin (hCS) mRNA and its expression
A;Reference number: I52412; MUID:83209060; PMID:6350050
A;Accession: I52412
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-<1> <JAN>
A;Cross-references: GB:M15415; NID:g337
R;Sherwood, L.M.; Bornstein, W.; Schechter, I.
Proc. Natl. Acad. Sci. U.S.A. 76, 6824-6828, 1979
A;Title: Primary structure of the NH2-terminal extracellular portion of human placental
A;Reference number: A93833; MUID:80054027; PMID:291015
A;Accession: A93833
A;Molecule type: protein
A;Residues: 1-<2> <SHE>
A;Experimental source: placenta
R;Shine, J.; Seeburg, P.H.; Martial, J.A.; Baxter, J.D.; Goodman, H.M.
Nature 270, 494-499, 1977
A;Title: Construction and analysis of recombinant DNA for human chorionic somatomammo
A;Reference number: A93812; MUID:78041751; PMID:925008
A;Accession: A93812
A;Molecule type: DNA
A;Residues: 50-217 <SHI>
A;Experimental source: placenta
R;Li, C.H.; Dixon, J.S.; Lyons, ...
Arch. Biochem. Biophys. 155, 95-101, 1973
A;Title: Amino acid sequence of human growth hormone-like
A;Reference number: A<...>; MUID:<...>; PMID:<...>
A;Accession: A<...>
A;Molecule type: protein
A;Residues: 27-217 <...>
A;Experimental source: placenta
R;Niall, H.D.
In: Isolation and characterization of [illegible] Workshop pp. [illegible]
A;Title: The chemistry of the human [illegible] hormones.
A;Reference number: A94427
A;Accession: A94427
A;Molecule type: protein
A;Residues: 27-217 <NIA>
A;Experimental source: placenta
R;Nir, A.; Sharir, N.; Lapidot, A.
Biochem. Soc. Trans. 1, 205, 1973
A;Title: Catecholamine methyltransferase from human placentae: purification and



[illegible]

LOCUS
DEFINITION p1L... 1 aa.
NID:Alternate names: growth hormone
OS:Species: Equus caballus (domestic horse)
[illegible] #Sequence_revision 11-Jul-[illegible] #text_change 27-Apr-[illegible]
AC:Accession: A91[illegible]; A91695; A91683; A90240; A91511
RA:Takio, M.M.; Poskus, E.; Ignatov, A.A.; Ferreira, F.; Santome, J.A.; Dellacha, J.M.; Pa
[illegible], J. Dept. Protein Res. 8, 435-444, 1976
AT:Title: Primary structure of equine growth hormone.
AR:Reference number: A91752; MUID:77005410; PMID:903751
AA:Accession: A91752
AM:Molecule type: protein
AR:Residues: 1-190 <ZAK>
RA:Takio, M.M.; Poskus, E.; Dellacha, J.M.; Paladini, A.C.; Santome, J.A.
[illegible], 34, 67-75, 1977
[illegible]: The amino acid sequence of equine growth hormone.
[illegible]Reference number: A91895; MUID:78208362; PMID:417849
AA:Accession: A91895
AM:Molecule type: protein
AR:Residues: 1-190 <ZAZ>
RA:Takio, M.M.; Poskus, E.; Dellacha, J.M.; Paladini, A.C.; Santome, J.A.
[illegible], 2[illegible], 1978
AT:Title: Amino acid sequences around the cystine residues in equine growth hormone.
AR:Reference number: A91683
AA:Accession: A91683
AM:Molecule type: protein
AR:Residues: 42-69;152-190 <ZAX>
RA:[illegible]; Harkiss, A.S.
[illegible], J. [illegible], 19-21, 1978
AT:Title: Amino acid sequences around the cystine residues in horse growth hormone.
AR:Reference number: A91024; MUID:[illegible]; PMID:[illegible]
AA:Accession: A91024
AM:Molecule type: protein
AR:Residues: 179-190 <ALL>
RA:[illegible]; Bromer W.
RL:Keywords: hormone; pituitary
FH:Key/Location/Qualifiers: 108/disulfide bonds: #status experimental

Query Match 91.4%; Score 159.5; DB 2; Length 190;
Best Local Similarity 68.1%; Pred. No. 1.5e-12;
Matches 42; Conservative 6; Mismatches 9; Indels 1; Gaps 1;

Qy . FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPTPSNREETQQKSNLELLRISLLLIQSWLEPVQFLRSVFANSLVYGASDSNVYDLLKDLEEGIQTLMGRLEDGSPRTGQIFKQTYSKFDTNSHNDDALLKNYGLLYCFRKDMDKVETFLRIVQCRSVEGSCGF
. FTAMFLSSLFANAVSAAVPLDQSVALVKTELEKAKFLKQKRSNDQNF 49

Search completed: [illegible]
Job time : [illegible] secs

[illegible]

LOCUS JK0219
DEFINITION African elephant
NID:Alternate names: growth hormone
OS:Species: Loxodonta africana (African elephant)
[illegible]Date: 03-Aug-1992 #Sequence_revision 03-Aug-1992 #text_change 15-Oct-1996
AC:Accession: JK0219
RA:Hobbs, R.S.; Rhodes, M.C.; Li, C.H.; Ham, S.C.L.
Int. J. Pept. Protein Res. 43, 468-472, 1989
AT:Title: Primary structure of elephant growth hormone.
AR:Reference number: JK0219
AA:Accession: JK0219
AM:Molecule type: protein
AR:Residues: 1-191 <HOB>
AT:Experimental source: pituitary gland
KW:Keywords: prolactin
FH:Key/Location/Qualifiers: #status experimental <MAT>

Query Match 91.4%; Score 159.5; DB 2; Length 191;
Best Local Similarity 68.1%; Pred. No. 1.5e-12;
Matches 42; Conservative 6; Mismatches 9; Indels 1; Gaps 1;

Searched: 112092 seqs, 41470428 residues

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

The best N.N is the number of sequences predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

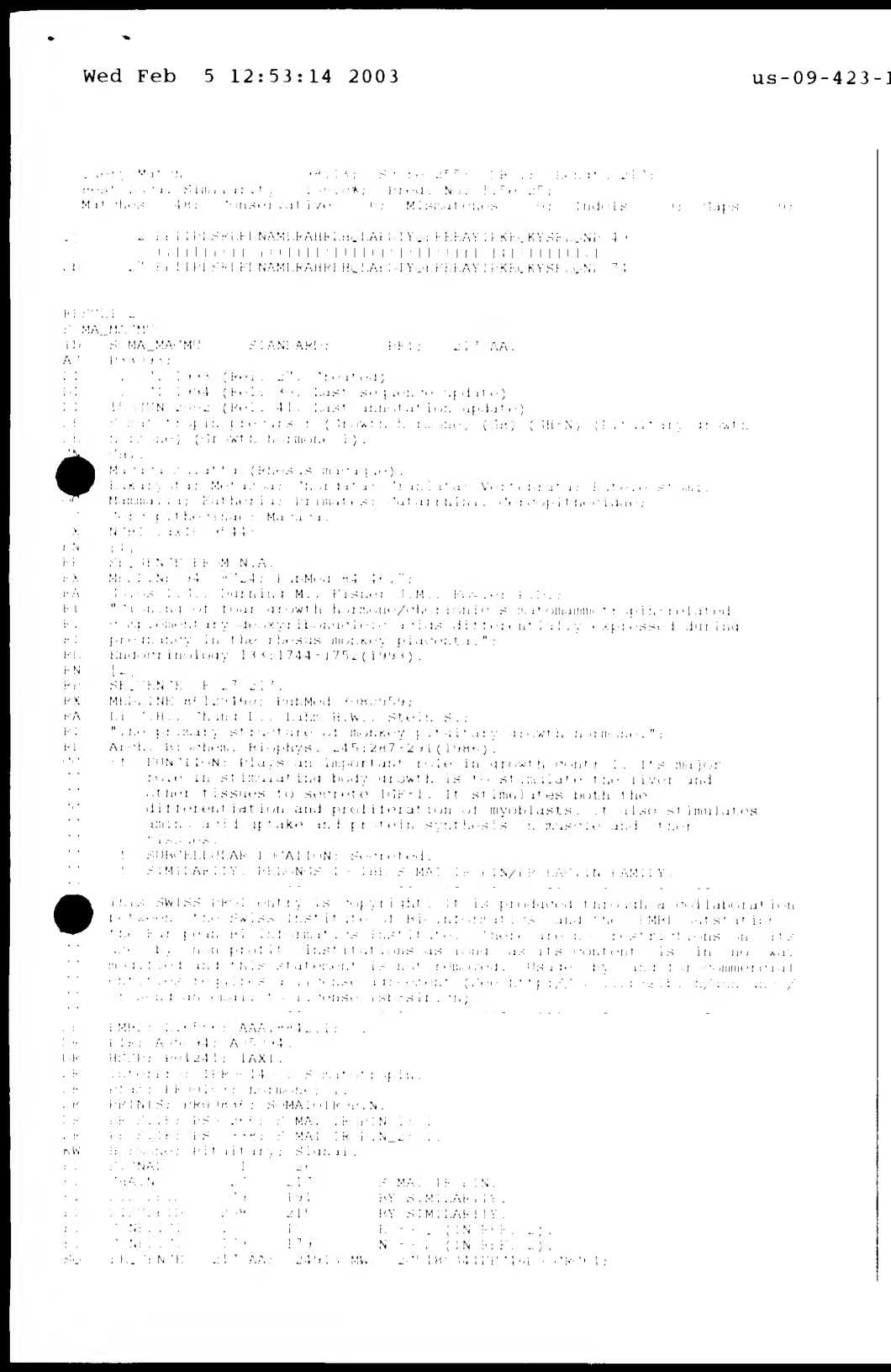
SUMMARIES

Result No.	Score	% Match	Length	DB	ID	Description
1	256	98.1	217	1	S4MA_HUMAN	P61981 Homo sapien
2	255	98.1	217	1	S4MA_MACMU	Q29185 Macaca mula
3	253	98.1	217	1	S4MA_PANTR	P54775 pan troglod
4	249	95.8	217	1	S4MA_CALJA	Q28085 callithrix
5	249	95.8	217	1	S4MA_SAIBB	P56343 saimiri bol
6	249	95.8	217	1	S4MA_PANTR	P54767 pan troglod
7	248	97.2	217	1	S4M2_HUMAN	P61142 homo sapien
8	194	76.6	217	1	S4M2_MACMU	Q28779 macaca mula
9	193	75.8	217	1	P61_HUMAN	P61143 homo sapien
10	172.5	62.1	217	1	S4MA_MUSA	Q07885 mus muscu
11	172.5	61.1	217	1	S4MA_RALB	P55092 balaenopter
12	172.5	61.1	217	1	S4MA_XENLA	P29392 xenopus la
13	172.5	61.1	217	1	S4MA_VITVI	Q12746 vulpes vul
14	172.5	61.1	217	1	S4MA_ANSA	P55711 anas banc
15	171.5	61.1	217	1	S4MA_DELLA	P49681 delphinap
16	171.5	61.1	217	1	S4MA_BOVSE	P61215 ovis aries
17	171.5	61.1	217	1	S4MA_MESAU	P55788 mesocricet
18	170.5	61.1	217	1	S4MA_ICT	P31218 ictalurus
19	169.5	61.1	216	1	S4MA_RABIT	P46407 oryctolagus
20	169.5	61.1	216	1	S4MA_RAT	P61141 rattus nor
21	169.5	61.1	217	1	S4MA_CAVPO	P49681 cavia porc
22	169.5	61.1	217	1	S4MA_MYOLU	Q29185 myotis luc
23	169.5	61.1	217	1	S4MA_MUSVI	Q29175 mustela vis
24	169.5	61.1	217	1	S4MA_LAMLA	Q07885 lama lama
25	169	60.8	217	1	S4MA_MELGA	P22772 meleagris g
26	148	56.9	191	1	S4MA_CHEMY	P51105 chelonia my
27	137	55.4	217	1	S4MA_M_NO	P49681 mobilpass
28	136	55.4	217	1	S4MA_TRUV	P49681 trichosurus
29	134	55.4	190	1	S4M1_ACTCH	P29774 actipenser
30	134	55.4	190	1	S4M1_ACIBA	P29774 acipenser
31	131	54.4	217	1	S4MA_BETCH	P49681 betula ver
32	131	54.4	217	1	S4MA_BOVIN	P61249 bos taurus
33	131.5	54.8	217	1	S4MA_CHICK	P56137 corvus corp
34	131.5	54.8	217	1	S4MA_MACMU	P49681 macaca mula
35	131.5	54.8	217	1	S4MA_PIG	P49681 pig
36	131	54.4	217	1	S4MA_RANCA	P49681 rana catesb
37	131	54.4	217	1	S4MA_RANPI	P49681 rana pipien
38	131	54.4	217	1	S4MA_RANTA	P49681 rana tagoia
39	131	54.4	217	1	S4MA_SHEEP	P49681 ovis aries
40	131	54.4	217	1	S4MA_XENBO	P49681 xenopus bor
41	131.5	54.8	217	1	S4MA_XENLA	P49681 xenopus lae
42	131	54.4	217	1	S4MA_XENTR	P49681 xenopus tro
43	130.5	54.4	217	1	S4MA_ANAPL	P49681 anas platyr
44	130.5	54.4	217	1	S4MA_PIGEO	P49681 columba liv
45	130	54.4	217	1	S4MA_XENMA	P49681 xenopus mul

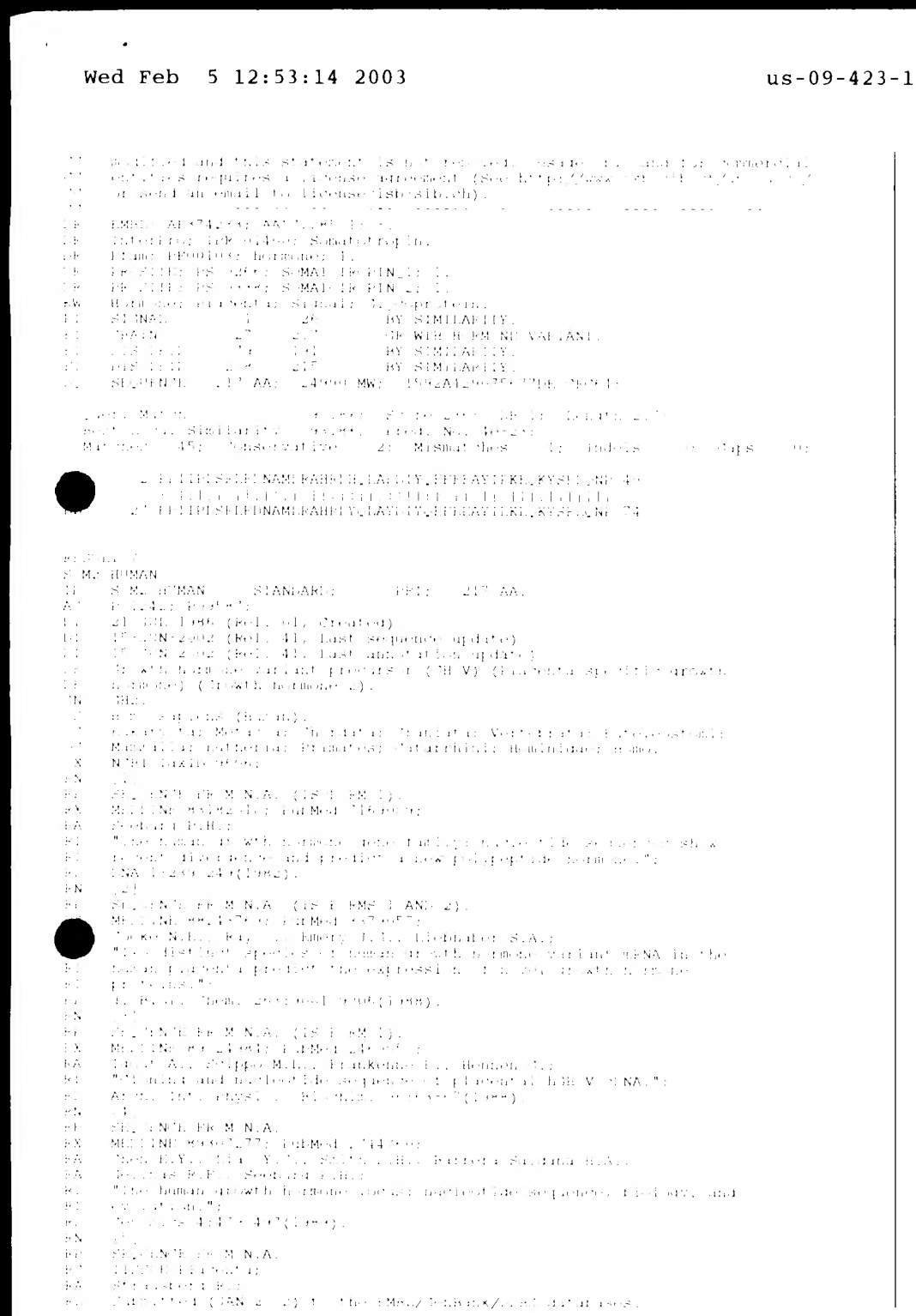
ALIGNMENTS

RESULT 1
S4MA_HUMAN
ID S4MA_HUMAN STANDARD; PRT; 217 AA.
AC P61981; P29374; P31948; Q13748; Q13749; Q16778; Q16779;
DT 21-JUL-1998 (Rel. 36, Created)
DT 01-MAR-2002 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Stathmin (Prosolin) (Oncoprotein 18) (Op18) (Leukemia-associated
DE hematopoietic gene protein) (Metablastin).
GN CHI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (CLONE HM...).
RX MEDLINE=90...; PubMed=...;
RA ...;
RA "...";
RT "...";
RL ...
RN [2]
RP SEQUENCE FROM N.A. (CLONE HM...).
RX MEDLINE=...; PubMed=...;
RA ...;
RT "...";
RL ...
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=...; PubMed=...;
RA ...;
RT "...";
RL ...
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=...; PubMed=...;
RA ...;
RT "...";
RL ...
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=...; PubMed=...;
RA ...;
RT "...";
RL ...
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=...; PubMed=...;
RA ...;
RT "...";
RL ...
RN [7]
RP SEQUENCE FROM N.A.
RX MEDLINE=...; PubMed=...;
RA ...;
RT "...";
RL ...
RN [8]
RP SEQUENCE FROM N.A.
RX MEDLINE=...; PubMed=...;
RA ...;
RT "...";
RL ...

```
RN   [7]
RC   SEQUENCE FROM N.A. (ISOFORM 4).
RA   [illegible];
RX   MEDLINE 20402671; PubMed 10931946;
RA   Hu R.-M., Han Z.-G., Song H.-D., Peng Y.-D., Huang Q.-H., Ren S.-X.,
RA   Xu S.-Y., Xu G.-F., Zhang C.-K., Yang H.-Y., Zhang J., Gu B.-W.,
RA   Tang W., Bai M., Mao Y.-F., Gao G.-F., Jiang C.-L., Jin B.-Q.,
RA   Shi J.-X., Chen Z., Chen M.-S., Chen J.-L.;
RT   "Gene expression profiling in the human hypothalamus-pituitary-adrenal
RT   axis and full-length cDNA cloning.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:9543-9548(2000).
RN   [8]
RP   SEQUENCE OF 1-20 FROM N.A.
RX   MEDLINE 86172565; PubMed 3013264;
RA   Gray C.J., Baldridge J.S., McKeown K.S., Heyneker H.L., Chang C.N.;
RT   "Periplasmic production of correctly processed human growth hormone in
RT   Escherichia coli: natural and bacterial signal sequences are
RT   interchangeable.";
RL   Gene 39:247-254(1985).
RN   [9]
RP   SEQUENCE OF 1-217.
RX   MEDLINE 82203292; PubMed 6812034;
RA   Li C.H., Dixon J.S., Liu W.-K.;
RT   "Human pituitary growth hormone. XIX. The primary structure of the
RT   hormone.";
RL   Arch. Biochem. Biophys. 213:444-450(1982).
RN   [10]
RP   SEQUENCE OF 24-217, AND REVISIONS.
RX   MEDLINE 76228292; PubMed 5813034;
RA   Li C.H., Dixon J.S., Liu W.-K.;
RT   "Human pituitary growth hormone. XIX. The primary structure of the
RT   hormone.";
RL   Arch. Biochem. Biophys. 169:444-450(1975).
RN   [11]
RP   SEQUENCE OF 24-217.
RX   MEDLINE 74158151; PubMed 4151151;
RA   Beaney J.A., Dixon J.S., Li C.H.;
RT   "Structural comparison of human pituitary growth hormone, human
RT   chorionic somatomammotropin, and ovine pituitary growth and
RT   lactogenic hormones.";
RL   Int. J. Pept. Protein Res. 6:281-287(1974).
RN   [12]
RP   SEQUENCE OF 37-61 AND 102-134.
RX   MEDLINE 71147765; PubMed 4270044;
RA   Niall H.D.;
RT   "Revised primary structure for human growth hormone.";
RL   Nat. New Biol. 230:90-91(1971).
RN   [13]
RP   REVISIONS TO 120 AND 152-159.
RX   MEDLINE 71145096; PubMed 5270529;
RA   Niall H.D., Hogan M.L., Sauer R., Rosenblum I.Y., Greenwood F.C.;
RT   "Sequences of pituitary and placental lactogenic and growth hormones:
RT   evolution from a primordial peptide by gene reduplication.";
RL   Proc. Natl. Acad. Sci. U.S.A. 68:866-869(1971).
RN   [14]
RP   REVIEW.
RA   Niall H.D.;
RT   "The chemistry of the human lactogenic hormones.";
RT   (In) Taylor S. (eds.);
RL   Proceedings of the Serono symposium: Advances in human growth hormone research,
RL   pp. 1-20, Alpha Omega Alpha Press, Cardiff (1972).
RN   [15]
RP   SEQUENCE OF 1-79 (ISOFORM 2).
RX   MEDLINE 91117661; PubMed 2002247;
RA   Chapman G.B., Rogers K.M., Rittenhouse J., Bourdeau H.A., Peters J.L.,
RA   Lipson K., Cary P.D., Tate Robinson C.;
RT   "The 20,000 molecular weight variant of human growth hormone.
RT   Preparation and some physical and chemical properties.";
RL   J. Biol. Chem. 265:4265-4270(1990).
RN   [16]
RP   SEQUENCE OF 16-80 (ISOFORM 2).
RX   MEDLINE [illegible], PubMed [illegible];
RA   [illegible];
RT   "[illegible]";
RL   [illegible].
RN   [17]
RP   SEQUENCE OF 1-191 (ISOFORM 3).
RA   Lewis U.J., Bergstrom R.C., Seavey B.K.;
RT   "The 22,000-dalton variant of human growth hormone.";
RL   Biochem. Biophys. Res. Commun. 92:1213-1219(1980).
RN   [18]
RP   CHARACTERIZATION OF GLN-65 AS ASN.
RX   MEDLINE 90252647; PubMed 2028470;
RA   Lewis U.J., Singh R.N., Seavey B.K.;
RT   "Alteration processing of human growth hormone at position 65.";
RT   deamidation.";
RL   J. Biol. Chem. 265:11243-11247(1990).
RN   [19]
RP   REVIEW.
RX   MEDLINE 74121145; PubMed 4275035;
RA   Isaksson O.;
RT   "Growth hormone heterogeneity in human pituitary and plasma.";
RL   Horm. Res. 51 Suppl. 1:2-5(1974).
RN   [20]
RP   3D-STRUCTURE MODELING.
RX   MEDLINE 94150077; PubMed 8247155;
RA   Cohen F.E., Kuntz I.D.;
RT   "Prediction of the three-dimensional structure of human growth
RT   hormone.";
RL   Proteins 17:62-76(1993).
RN   [21]
RP   X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX   MEDLINE 92187547; PubMed 1540774;
RA   de Vos A.M., Ultsch M., Kossiakoff A.A.;
RT   "Human growth hormone and extracellular domain of its receptor:
RT   crystal structure of the complex.";
RL   Science 255:306-312(1992).
RN   [22]
RP   X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
RX   MEDLINE 95037540; PubMed 7984144;
RA   Somers W., Ultsch M., de Vos A.M., Kossiakoff A.A.;
RT   "The X-ray structure of a growth hormone-prolactin receptor complex.";
RL   Nature 372:478-481(1994).
RN   [23]
RP   X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
RA   Chantalat L., Chirgadze N.Y., Jones N., Korber F., Nakada L.,
RA   Nagel G.A., Wlodawer A.;
RT   "The crystal structure of wild-type human growth hormone at 2.7 A
RT   resolution.";
RL   Protein Pept. Lett. 2:337-346(1995).
RN   [24]
RP   X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX   MEDLINE 97119054; PubMed 8943044;
RA   Sundstrom M., Lundqvist L., Bowtin L., Lobez-Fabo M., Wikstrom ...
RA   Norstedt G.;
RT   "Crystal structure of an antagonist mutant of human growth ...
RT   alone, in complex with its receptor, and ... A resolution.";
RL   J. Biol. Chem. 271:32107-32110(1996).
CC   -!- FUNCTION: Plays an important role in growth control. It is a
CC       major stimulator of overall body growth. It also stimulates
CC       other tissues to secrete IGF-1. It stimulates both the
CC       differentiation and proliferation of chondrocytes. It also
CC       enhances amino acid uptake and protein synthesis in muscle and
CC       tissues.
CC   -!- SUBUNIT: Monomer. Binds 1 zinc and possibly 2 zinc ions per
CC       molecule. Exists as monomeric, dimeric, trimeric and tetrameric
CC       heteropolymeric combinations. Can also form a complex with
CC       GHBP or with the alpha-subunit of the ...
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- ALTERNATIVE PRODUCTS: A least 4 isoforms; 22 kDa (the main form),
CC       20 kDa variant, 6 and 8 are not known to be ...
CC       ...ficient.
CC   -!- DISEASE: DEFECTS IN GHRL ARE A CAUSE OF PITUITARY DWARFISM I AND
CC       IV.
CC   -!- PHARMACEUTICAL: Available under the names Nutropin (Genentech), Protropin (Genentech), Humatrope (Eli Lilly) and Saizen or Serostim (Serono). Used for treatment of growth hormone deficiency and of Turner's syndrome.
```



```
ID   [illegible] STANDARD; PRT; [illegible] AA.
AC   P[illegible];
DT   15-JUN-[illegible] (Rel. 41, Created)
DT   15-JUN-[illegible] (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Growth hormone variant precursor (GH-V) (Chorionic somatomammotropin
DE   hormone) (Growth hormone).
GN   GH2.
OS   Pan troglodytes (Chimpanzee).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX   NCBI_TaxID=9598;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Fox J.A., Espiritu C.L., Sanchez C., Barrera Saldana H.;
RT   "Independent duplication of the growth hormone gene in three
RT   Anthropoidean lineages.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Plays an important role in growth control. Its major
CC       role in stimulating body growth is to stimulate the liver and
CC       other tissues to secrete IGF-1. It stimulates both the
CC       differentiation and proliferation of myoblasts. It also stimulates
CC       amino acid uptake and protein synthesis in muscle and other
CC       tissues.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- [illegible] SPECIFICITY: [illegible] based on its presence.
CC   -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; AF[illegible]; AAK[illegible].1   -.
DR   InterPro; IPR[illegible]; Somatotropin.
DR   Pfam; PF00103; hormone; 1.
DR   PROSITE; PS00266; SOMATOTROPIN_1; 1.
DR   PROSITE; PS00338; SOMATOTROPIN_2; 1.
KW   Hormone; Pituitary; Signal.
FT   SIGNAL       1     26       BY SIMILARITY.
FT   CHAIN       27    217       SOMATOTROPIN.
FT   DISULFID    79    191       BY SIMILARITY.
FT   DISULFID    [illegible]     BY SIMILARITY.
SQ   SEQUENCE   217 AA;  24[illegible] MW;  [illegible]CRC64;

  Query Match          [illegible]%;  Score 244;  DB 1;  Length 217;
  Best Local Similarity   97.9%;  Pred. No. 8.7e-25;
  Matches   47;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEAAYIPKEQKYSFLQNP 49
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEAAYIPKEQKYSFLQNP 48


RESULT 5
S_MA_SAIBO
ID   S_MA_SAIBO     STANDARD;      PRT:  217 AA.
AC   P[illegible];
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Somatotropin precursor (Growth hormone).
GN   GH.
OS   Saimiri boliviensis boliviensis (Bolivian squirrel monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae; Saimiri.
OX   NCBI_TaxID=39432;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21[illegible];
RA   [illegible] M.M., Makova K.D., Adkins R.M., Chen S., Li W.H.;
RT   "Episodic evolution of growth hormone in primates and emergence of the
RT   [illegible] specificity of human growth hormone receptor.";
```



```
AC   xxxxxx; xxxxxx;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUN-2000 (Rel. 41, Last annotation update)
DE   Growth hormone variant precursor (GH-V) (Placental growth hormone, GH-2)
DE   (Growth hormone-2).
GN   GH2.
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OX   NCBI_TaxID=9544;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Golos T.G.;
RL   Submitted (JUN-1994) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=94304624; PubMed=8194617;
RA   Seely J.C., Dutting M., Fisher J.M., Fowler T.J.;
RT   "Cloning of four growth hormone/chorionic somatomammotropin-related
RT   messenger deoxyribonucleic acids differentially expressed during
RT   pregnancy in the rhesus monkey (Macaca mulatta).";
RL   Endocrinology 134:1744-1752(1994).
CC   -!- FUNCTION: Plays an important role in or with control. Its major
CC       role in stimulation body growth is to stimulate the liver and
CC       other tissues to secrete IGF-1. It stimulates both the
CC       differentiation and proliferation of myoblasts. It also stimulates
CC       amino acid uptake and protein synthesis in muscle and other
CC       tissues.
CC   -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC   -!- TISSUE SPECIFICITY: Expressed in the placenta.
CC   -!- SIMILARITY: BELONGS TO THE GROWTH HORMONE FAMILY.
CC   -!- ----------- ---------- ---------- ---------- ---------- ----------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------- ---------- ---------- ---------- ---------- ----------
DR   EMBL; L22464; AAA30804.1; -.
DR   EMBL; L35055; AAA20804.1; -.
DR   HSSI; I60124; IHHU.
DR   InterPro; IPR001400; Somatotropin.
DR   Pfam; PF00103; hormone; 1.
DR   PRINTS; PR00836; SOMATOTROPIN.
DR   to SIB; isb-2xx; SOMATOTROPIN_1; 1.
DR   to SIB; isb-xxx; SOMATOTROPIN_2; 1.
KW   Hormone; Placenta; Signal; Glycoprotein.
FT   SIGNAL       1     26       BY SIMILARITY.
FT   CHAIN       27    217       GROWTH HORMONE VARIANT.
FT   DISULFID    53    191       BY SIMILARITY.
FT   DISULFID   208    215       BY SIMILARITY.
FT   CONFLICT    57     57       L -> F (IN REF. 2).
FT   CONFLICT   152    152       F -> S (IN REF. 2).
SQ   SEQUENCE   217 AA;  25221 MW;  8C61CF8B9C24EA00 CRC64;

  Query Match             %;  Score 1xx;  DB 2;  Length 217;
  Best Local Similarity   %;  Pred. No. 2.2e-1x;
  Matches   xx;  Conservative   x;  Mismatches   x;  Indels   0;  Gaps   0;

Qy       1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQ 50
           ||||  |  | | || |||||||| | |||||||||||||||||||||| |
Db     27 MFPTIPLSRLFDNAVLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQ 76

RESULT 2
...
ID   xxxxx_MACMU     STANDARD;      PRT;   217 AA.
AC   xxxxxx;
```

```
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-APR-1990 (Rel. 17, Last sequence update)
DT   15-JUN-2000 (Rel. 41, Last annotation update)
DE   Lactogen precursor (Chorionmammotropin, Chorionic somatomammotropin) (CS).
GN   CSH1 AND CSH2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (GENE: CSH1).
RX   MEDLINE=85051126; PubMed=6208952;
RA   Selby M.J., Barta A., Boorstein W.R., Richards R.I., Eberhardt N.L.;
RT   "Analysis of a major human chorionic somatomammotropin gene and two
RT   for two functional promoter elements.";
RL   J. Biol. Chem. 259:14758-14765(1984).
RN   [2]
RP   SEQUENCE FROM N.A. (GENE: CSH1).
RX   MEDLINE=87161295; PubMed=3468695;
RA   Eberhardt N.L., Kim S.-Y., Bitzer M.J., Yen P.Y., Seeburg P.H.,
RA   Eberhardt N.L., Barta A.;
RT   "The human growth hormone gene locus: structure, evolution, and
RT   allelic variations.";
RL   DNA 6:59-70(1987).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87160216; PubMed=3468694;
RA   Barrera-Saldana H.A., Seeburg P.H., Saunders G.F.;
RT   "Two structurally different genes produce the same secreted hormone
RT   placentae lactogen.";
RL   J. Biol. Chem. 262:8787-8793(1987).
RN   [4]
RP   SEQUENCE FROM N.A. (GENES: CSH1 AND CSH2).
RX   MEDLINE=20147477; PubMed=10711355;
RA   Chen E.Y., Liao Y.C., Smith D.H., Barrera-Saldana H.A., Gelinas R.E.,
RA   Seeburg P.H.;
RT   "The human growth hormone gene family: nonsense mutations and
RT   evolution.";
RL   Genomics 4:479-497(1989).
RN   [5]
RP   SEQUENCE.
RX   MEDLINE=84186210; PubMed=6698980;
RA   Seeburg P.H.;
RT   "The human growth hormone gene family: structure and evolution of
RT   recent divergence and predict a new polypeptide hormone.";
RL   DNA 1:239-249(1982).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RA   Seeburg P.H.;
RL   Submitted (JAN-1982) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   SEQUENCE OF 27-217 FROM N.A.
RX   MEDLINE=79226019; PubMed=37...;
RA   Shine J., Seeburg P.H., Martial J.A., Baxter J.D., Goodman H.M.;
RT   "Construction and analysis of recombinant DNA for human chorionic
RT   somatomammotropin.";
RL   Nature 270:494-499(1977).
RN   [8]
RP   SEQUENCE OF 27-217.
RX   MEDLINE=...; PubMed=...;
RA   Li C.H., Dixon J.S., Chung D.;
RT   "Amino acid sequence of human chorionic somatomammotropin.";
RL   Arch. Biochem. Biophys. ...:... (19..).
RN   [9]
RP   SEQUENCE OF 27-217.
RX   MEDLINE=...; PubMed=...;
RA   Shirai T.M., Barrera-Saldana H., Morgan W., ..., Talmadge M.;
RT   "Amino acid sequence of human placental lactogen.";
RL   Nature New Biol. ...:... (19..).
RN   [10]
RP   LINKAGE.
RA   Shirai T.M., Barrera-Saldana H., Morgan W.L., Talmadge M.;
```

```
EL   Nature New Biol. 245:64-66(1973).
RN   [1]
RA   SWISS-PROT; DATA BANK.
RX   MEDLINE; 74057015.
RA   Bairoch A., Boeckmann B., Ferro S., Gasteiger E.,
RT   "Identification of the interchain disulfide bonds of dimeric human
RT   placental lactogen.";
RL   Submitted (JUL-1989) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- MISCELLANEOUS: THE SEQUENCE SHOWN IS KNOWN.
CC   -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyrighted. It is made available under the
CC   terms of the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; X00121; CAA24484.1;  -.
DR   EMBL; J00265; AAA40747.1;  -.
DR   EMBL; K01316; AAA52139.1;  -.
DR   EMBL; M14120; AAA52551.1;  -.
DR   EMBL; J00756; AAA40921.1;  -.
DR   EMBL; M59958; AAA52121.1;  -.
DR   EMBL; M20074; AAA40755.1;  -.
DR   PIR; A25445; GHBP.
DR   PIR; A26445; A26445.
DR   PIR; S32435; S32435.
DR   PIR; E32435; E32435.
DR   HSSP; P01241; 1A22.
DR   Genew; HGNC:2441; SHGH.
DR   Genew; HGNC:2444; SHGH1.
DR   MIM; 150200; -.
DR   InterPro; IPR001400; Somatotrophin.
DR   Pfam; PF00103; Somatotrophin.
DR   PRINTS; PR00837; SOMATOTROPIN.
DR   PROSITE; PS00264; SOMATOTROPIN_1; 1.
DR   PROSITE; PS00338; SOMATOTROPIN_2; 1.
KW   Hormone; Disulfide bond; Multigene family; Signal.
FT   SIGNAL        1     26
FT   CHAIN        27    217       GROWTH HORMONE.
FT   DISULFID     53    165       INTERCHAIN (WITH P01241).
FT   DISULFID     58    182       INTERCHAIN (WITH P01241).
FT   DISULFID     56     56       INTERCHAIN (WITH P01241).
FT   VARIANT      54     54       A -> V (IN GHB).
FT                                /FTId=VAR_000202.
FT   VARIANT      64     64       S -> L (IN GHB).
FT                                /FTId=VAR_000203.
FT   CONFLICT     91     91       I -> T (IN REF. 4).
FT   CONFLICT     97     97       MISSING (IN REF. 3).
FT   CONFLICT    119    119       MISSING (IN REF. 3).
FT   CONFLICT    144    145       SD -> RS (IN REF. 3).
SQ   SEQUENCE   217 AA;  25020 MW;  2A5B0D07A7184141 CRC64;

  Query Match             95.4%;  Score 1417;  DB 1;  Length 217;
  Best Local Similarity   99.5%;  Pred. No. 1.1e-150;
  Matches 216;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPTP
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPTP

Qy      61 SNREETQQKSNLELLRISLLLIQSWLEPVQFLRSVFANSLVYGASDSNVYDLLKDLEEGIQ
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 SNREETQQKSNLELLRISLLLIQSWLEPVQFLRSVFANSLVYGASDSNVYDLLKDLEEGIQ

Qy     121 TLMGRLEDGSPRTGQIFKQTYSKFDTNSHNDDALLKNYGLLYCFRKDMDKVETFLRIVQCR
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TLMGRLEDGSPRTGQIFKQTYSKFDTNSHNDDALLKNYGLLYCFRKDMDKVETFLRIVQCR

Qy     181 SVEGSCGF
           ||||||||
Db     181 SVEGSCGF

RESULT 2
SOMA_RAT
ID   SOMA_RAT       STANDARD;      PRT;   217 AA.
AC   P04631;
DT   01-JAN-1988 (Rel. 07, Created)
DT   15-JUN-2000 (Rel. 38, Last sequence update)
DT   15-JUN-2000 (Rel. 38, Last annotation update)
DE   Somatotropin precursor (Growth hormone).
GN   GH1 OR GH.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 87120006; PubMed=3027677;
RA   Cooke N.E., Baxter J.D.;
RT   "Structure of the rat growth hormone gene family: parallel
RT   evolution of primate and rodent growth hormones.";
RL   Endocr. Rev. 7:91-116(1986).
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyrighted. It is made available under the
CC   terms of the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; X06369; CAA29611.1;  -.
DR   EMBL; Z46667; CAA86711.1;  -.
DR   PIR; B23911; SIRT.
DR   HSSP; P01241; 1A22.
DR   MGD; MGI:97579; Gh.
DR   GO; GO:0005179; F:hormone activity; IDA.
DR   Pfam; PF00103; Somatotrophin; 1.
DR   PRINTS; PR00837; SOMATOTROPIN.
DR   PROSITE; PS00264; SOMATOTROPIN_1; 1.
DR   PROSITE; PS00338; SOMATOTROPIN_2; 1.
KW   Hormone; Signal.
FT   SIGNAL        1     26       BY SIMILARITY.
FT   CHAIN        27    217       SOMATOTROPIN.
FT   DISULFID     58    182       BY SIMILARITY.
FT   DISULFID     62     62       BY SIMILARITY.
SQ   SEQUENCE   217 AA;  25102 MW;  [illegible] CRC64;

  Query Match             [illegible]  Score [illegible];  DB 1;  Length 217;
  Best Local Similarity   [illegible]  Pred. No. [illegible];
  Matches [illegible];  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPTP
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPTP

Qy      61 SNREETQQKSNLELLRISLLLIQSWLEPVQFLRSVFANSLVYGASDSNVYDLLKDLEEGIQ
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 SNREETQQKSNLELLRISLLLIQSWLEPVQFLRSVFANSLVYGASDSNVYDLLKDLEEGIQ

Qy     121 TLMGRLEDGSPRTGQIFKQTYSKFDTNSHNDDALLKNYGLLYCFRKDMDKVETFLRIVQCR
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TLMGRLEDGSPRTGQIFKQTYSKFDTNSHNDDALLKNYGLLYCFRKDMDKVETFLRIVQCR

Qy     181 SVEGSCGF
           ||||||||
Db     181 SVEGSCGF

RESULT 3
SOMA_BALMU
ID   SOMA_BALMU     STANDARD;      PRT;   217 AA.
AC   P04632;
DT   01-JAN-1988 (Rel. 07, Created)
DT   [illegible] (Rel. [illegible], Last sequence update)
DT   15-JUN-2000 (Rel. 38, Last annotation update)
```

```
ID   Somatotropin (Growth hormone).
AN   Sec.
XX   Molecule given to results (50 chars).
XX   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
XX   Mammalia; Eutheria; Cetartiodactyla; Cetacea; Mysticeti;
XX   Balaenidae; Balaenoptera.
XX   NCBI_TaxID=9767;
RN   [1]
RC   SEQUENCE.
RX   MEDLINE=85[illegible]; PubMed=2115[illegible];
RA   Tsuboi N.A., Tomoko T.A., Sugato A.A., Shipova T.A.;
RT   "Amino acid sequence of soiwhale somatotropin.";
RL   Int. Kolmogorov 4':1059-1[illegible](1992).
RN   [2]
RC   PRELIMINARY PARTIAL SEQUENCE.
RA   Shipova T.A., Bolatov A.A., Ponkov Y.A.;
RT   "Structural studies of tryptic peptides from large-granular bromide
RT   fragments of soiwhale (Balaenoptera borealis) somatotropin.";
RL   Biorgan. Khim. 4:1587-1597(1978).
CC   -!- FUNCTION: Plays an important role in growth control. Its major
CC       role in stimulating body growth is to stimulate the liver and
CC       other tissues to secrete IGF-1. It stimulates both the
CC       differentiation and proliferation of myoblasts. It also stimulates
CC       amino acid uptake and protein synthesis in muscle and other
CC       tissues.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
DR   EMBL; JN0487; JN0487.
DR   HSSP; P01241; 1AXI.
DR   InterPro; IPR001400; Somatotropin.
DR   Pfam; PF00103; 1.
DR   PRINTS; PR00836; SOMATOTROPIN.
DR   ProSite; PS00266; SOMATOTROPIN_1; 1.
DR   ProSite; PS00267; SOMATOTROPIN_2; 1.
KW   Hormone; Pituitary.
FT   DISULFID     52     184       BY SIMILARITY.
FT   DISULFID     180    199       BY SIMILARITY.
SQ   SEQUENCE   190 AA;  21761 MW;  [illegible] SOMATOTROPIN.

  Query Match             61.7%;  Score 750.5;  DB 1;  Length 190;
  Best Local Similarity   98.1%;  Pred. No. 1.6e-54;
  Matches  [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy      [illegible]
Db      [illegible]

RESULT 2
SOMA_VULVU
ID   SOMA_VULVU     STANDARD;      PRT;   190 AA.
AC   P[illegible];
DT   [illegible] (Rel. 11, Created)
DT   [illegible] (Rel. 11, Last sequence update)
DT   15-JUN-2004 (Rel. 41, Last annotation update)
DE   Somatotropin (Growth hormone).
GN   GH1.
OS   Vulpes vulpes (Red fox).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Vulpes.
OX   NCBI_TaxID=9627;
RN   [1]
RC   SEQUENCE.
RC   TISSUE=Pituitary;
RX   MEDLINE=87254270; PubMed=2722401;
RA   Li C.H., Dzhobski J., Chung D.;
RT   "Primary structure of fox pituitary growth hormone.";
RL   Int. J. Pept. Protein Res. 29:78-72(1987).
CC   -!- FUNCTION: Plays an important role in growth control. Its major
CC       role in stimulating body growth is to stimulate the liver and
CC       other tissues to secrete IGF-1. It stimulates both the
CC       differentiation and proliferation of myoblasts. It also stimulates
CC       amino acid uptake and protein synthesis in muscle and other
CC       tissues.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
DR   EMBL; U54120; AAB42[illegible].
DR   HSSP; P[illegible]; 1AXI.
DR   InterPro; IPR001400; Somatotropin.
DR   Pfam; PF00103; 1.
DR   PRINTS; PR00836; SOMATOTROPIN.
DR   ProSite; PS00266; SOMATOTROPIN_1; 1.
DR   ProSite; PS00267; SOMATOTROPIN_2; 1.
KW   Hormone; Pituitary.
FT   DISULFID     52     184       BY SIMILARITY.
FT   DISULFID     180    199       BY SIMILARITY.
SQ   SEQUENCE   190 AA;  21761 MW;  [illegible] SOMATOTROPIN.

  Query Match             61.7%;  Score 750.5;  DB 1;  Length 190;
  Best Local Similarity   98.1%;  Pred. No. 1.6e-54;
  Matches  [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy      [illegible]
Db      [illegible]

RESULT 3
SOMA_LOXAF
ID   SOMA_LOXAF     STANDARD;      PRT;   190 AA.
AC   P[illegible];
DT   [illegible] (Rel. 17, Created)
DT   [illegible] (Rel. 17, Last sequence update)
DT   [illegible] (Rel. 17, Last annotation update)
DE   Somatotropin (Growth hormone).
GN   GH1.
OS   Loxodonta africana (African elephant).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Proboscidea; Elephantidae; Loxodonta.
OX   NCBI_TaxID=9785;
RN   [1]
RC   SEQUENCE.
RA   Holmes J.D., Minder M.A., Li C.H., Pan Y.C.E.;
RT   "Primary structure of elephant growth hormone.";
RL   Int. J. Pept. Protein Res. 33:69-87(1989).
CC   -!- FUNCTION: Plays an important role in growth control. Its major
CC       role in stimulating body growth is to stimulate the liver and
CC       other tissues to secrete IGF-1. It stimulates both the
CC       differentiation and proliferation of myoblasts. It also stimulates
CC       amino acid uptake and protein synthesis in muscle and other
CC       tissues.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
DR   EMBL; U54120; AAB42[illegible].
DR   HSSP; P[illegible]; 1AXI.
DR   InterPro; IPR001400; Somatotropin.
DR   Pfam; PF00103; 1.
DR   PRINTS; PR00836; SOMATOTROPIN.
DR   ProSite; PS00266; SOMATOTROPIN_1; 1.
DR   ProSite; PS00267; SOMATOTROPIN_2; 1.
KW   Hormone; Pituitary.
FT   DISULFID     52     184       BY SIMILARITY.
FT   DISULFID     180    199       BY SIMILARITY.
SQ   SEQUENCE   190 AA;  21761 MW;  [illegible] SOMATOTROPIN.

  Query Match             61.7%;  Score 750.5;  DB 1;  Length 190;
  Best Local Similarity   98.1%;  Pred. No. 1.6e-54;
  Matches  [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy      [illegible]
Db      [illegible]

RESULT 4
```

```
[illegible]MA[illegible]
ID   [illegible]MA[illegible]    STANDARD;    PRT;   [illegible] AA.
AC   [illegible]
DT   [illegible] (Rel. [illegible], Created)
DT   [illegible] (Rel. [illegible], Last sequence update)
DT   [illegible] (Rel. [illegible], Last annotation update)
DE   [illegible] (GH) (Growth hormone).
GN   [illegible]
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible] Martinez [illegible], Barrera Saldana H.A.;
RT   "A [illegible] growth hormone cDNA codes for a mature protein identical to
RT   [illegible] growth hormone.";
RL   Gene [illegible](1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   van Leeuwen [illegible], Teske [illegible], van Garderen [illegible], Rutteman [illegible], M[illegible];
RT   "Extrapituitary growth hormone expression in the dog is initiated at
RT   the normal pituitary transcription start site in the mammary gland and
RT   at multiple upstream sites in lymphoid cells.";
RL   Submitted ([illegible]) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   [illegible] Mammary gland;
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible] van Leeuwen [illegible];
RT   "[illegible] mammary growth hormone gene transcription initiates at the
RT   pituitary specific start site in the absence of [illegible].";
RL   Mol. Cell. Endocrinol. [illegible](1999).
CC   -!- FUNCTION: Plays an important role in growth control. Its major
CC       role in stimulating body growth is to stimulate the liver and
CC       other tissues to secrete IGF-1. It stimulates both the
CC       differentiation and proliferation of myoblasts. It also stimulates
CC       amino acid uptake and protein synthesis in muscle and other
CC       tissues.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://[illegible]
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; [illegible]; AA[illegible]; -.
DR   EMBL; [illegible]; AAF[illegible]; -.
DR   EMBL; AF[illegible]; AAD[illegible]; -.
DR   PIR; [illegible]; S[illegible].
DR   HSSP; P01246; 1BST.
DR   InterPro; IPR001400; Somatotropin.
DR   Pfam; PF00103; hormone; 1.
DR   PRINTS; PR00836; SOMATOTROPIN.
DR   PROSITE; PS00266; SOMATOTROPIN_1; 1.
DR   PROSITE; PS00338; SOMATOTROPIN_2; 1.
KW   Hormone; Pituitary; Signal.
FT   SIGNAL        1     26       BY SIMILARITY.
FT   CHAIN        27    216       SOMATOTROPIN.
FT   DISULFID     79    189       BY SIMILARITY.
FT   DISULFID    206    214       BY SIMILARITY.
FT   CONFLICT      4      4       [illegible] (IN REF. 1).
FT   CONFLICT      7      7       N -> [illegible] (IN REF. 1).
SQ   SEQUENCE   216 AA;  24460 MW;  A8AD[illegible] CRC64;

  Query Match             [illegible];  Score [illegible];  DB 1;  Length 216;
  Best Local Similarity   68.1%;  Pred. No. [illegible];
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];
```

```
RT   [illegible]
RL   [illegible]

[illegible]
SOMA_[illegible]
ID   SOMA_[illegible]    STANDARD;    PRT;   [illegible] AA.
AC   [illegible]
DT   [illegible] (Rel. [illegible], Created)
DT   [illegible] (Rel. [illegible], Last sequence update)
DT   [illegible] (Rel. [illegible], Last annotation update)
DE   Somatotropin precursor (GH) [illegible].
GN   GH1.
OS   Felis silvestris [illegible].
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; [illegible].
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pituitary;
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Warren W.[illegible];
RT   "[illegible]";
RL   Gene [illegible].
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=[illegible];
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible];
RT   "[illegible]";
RL   Gene [illegible].
CC   -!- FUNCTION: Plays an important role in growth control. [illegible]
CC       [illegible]
CC       other tissues to secrete [illegible]
CC       differentiation and proliferation [illegible]
CC       amino acid uptake and protein synthesis [illegible]
CC       tissues.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://[illegible]
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; [illegible]; AA[illegible]; -.
DR   EMBL; [illegible]; AA[illegible]; -.
DR   HSSP; P01246; 1BST.
DR   InterPro; IPR001400; Somatotropin.
DR   Pfam; PF00103; hormone; 1.
DR   PRINTS; PR00836; SOMATOTROPIN.
DR   PROSITE; PS00266; SOMATOTROPIN_1; 1.
DR   PROSITE; PS00338; SOMATOTROPIN_2; 1.
KW   Hormone; Pituitary; Signal.
FT   SIGNAL        1     26       BY SIMILARITY.
FT   CHAIN        27    [illegible]       SOMATOTROPIN.
FT   DISULFID     79    [illegible]       BY SIMILARITY.
FT   DISULFID    [illegible]    [illegible]       BY SIMILARITY.
FT   CONFLICT    [illegible]
FT   CONFLICT    [illegible]
FT   CONFLICT    [illegible]
FT   CONFLICT    [illegible]
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             [illegible];  Score [illegible];  DB 1;  Length [illegible];
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];
```

Application: protein search, sw model

Run on: February 5, 2003, 12:51:21 ; Search time 2.04 Seconds
(without alignments)
[illegible] Million cell updates/sec

Title: US-09-423-100-1
Perfect score: 260
Sequence: 1 MSTTPLSPLPTNAMKAHP.....IEEAYTEEQKSSCL_NP 19

Scoring table: BLOSUM62
Gapop [illegible] Gapext [illegible]

Searched: 67180 seqs, 2 [illegible]47119 residues

Total number of hits satisfying chosen parameters: [illegible]

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_rodent:*
8: sp_short:*
9: sp_synthetic:*
10: sp_unclassified:*
11: sp_virus:*
12: sp_vertebrate:*
13: sp_unclassified:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_vertebrate:*
17: sp_archaea:*

Post-P Q is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

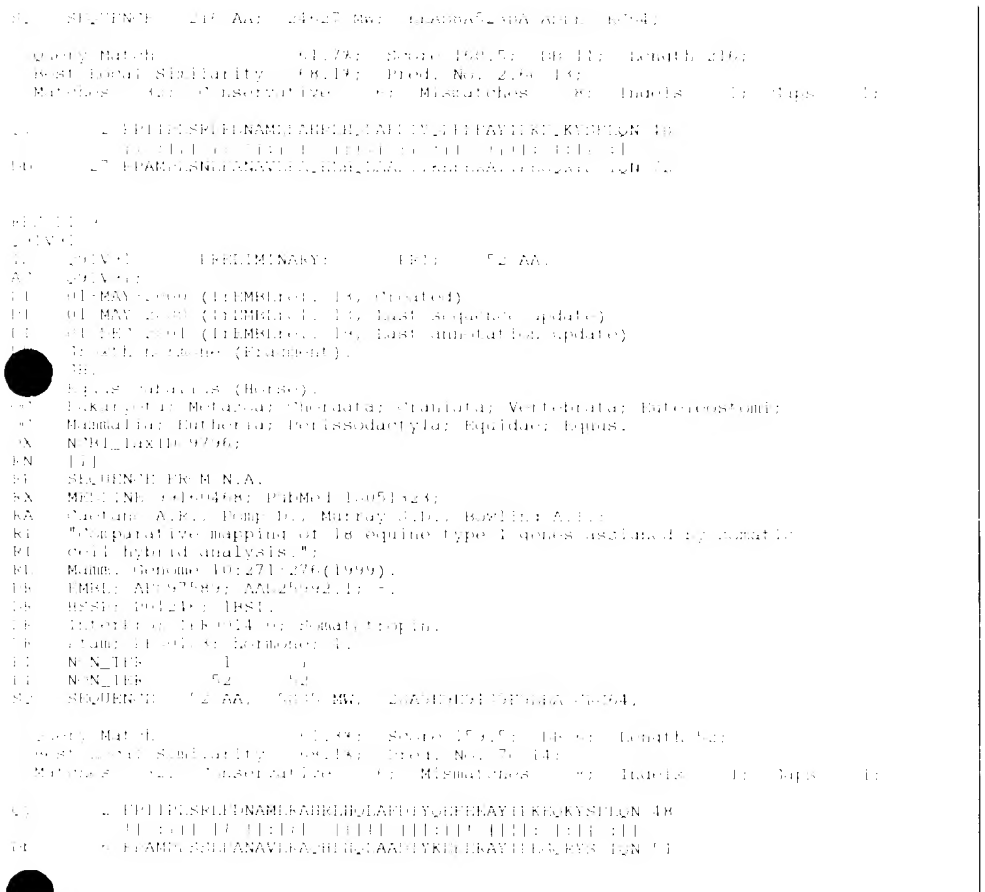
Result [illegible]
No. Score Match Length DB ID Description
[illegible]

ALIGNMENTS

RESULT 1
[illegible]
ID [illegible]
AC [illegible]
DT [illegible]
DT [illegible]
DT [illegible]
DE [illegible]
GN [illegible]
OS [illegible]
OC [illegible]
OC [illegible]
RN [1]
RP [illegible]
RA [illegible]
RT "Independent [illegible]
RL [illegible]
RL [illegible]
DR [illegible]
DR [illegible]
DR [illegible]
DR [illegible]
DR [illegible]
SQ [illegible]

Matches [illegible]

RESULT 2
[illegible]
ID [illegible]



```
RL   J. Clin. Endocrinol. Metab. 82:2[illegible] 1997 (online).
RA   Ma...
DR   HSSP; P01241; 1A22.
DR   InterPro; IPR001050; Somatotropin.
DR   Pfam; PF00103; hormone; 1.
DR   PRINTS; PR00837; SOMATOTROPIN.
DR   PROSITE; PS00264; SOMATOTROPIN_1; 1.
DR   PROSITE; PS00266; SOMATOTROPIN_2; 1.
KW   Signal.
FT   SIGNAL       1     26       POTENTIAL.
SQ   SEQUENCE   217 AA;  24[illegible] MW;  [illegible] CRC64;

  Query Match             56.5%;  Score 152;  DB 6;  Length 217;
  Best Local Similarity   [illegible]%;  Pred. No. [illegible];
  Matches  [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

QY       2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYS 44
           ||||||||||| |||||  ||||| ||||||  |||||||||||
Db       4 FPSVPLSRLFANAVLRAQHLHQLAADTYKEFERIYIPEGQRYS 48


RESULT 14
Q9Y5M6
ID   Q9Y5M6       PRELIMINARY;      PRT;   204 AA.
AC   Q9Y5M6;
DT   01-OCT-2000 (TrEMBLrel. 16, Created)
DT   01-OCT-2000 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Growth hormone (Fragment).
OS   Tursiops aduncus (Indian).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Odontoceti; Delphinidae; Tursiops.
OX   NCBI_TaxID=9739;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22065[illegible]; PubMed=[illegible];
RA   Liu J.Y., Makova K.D., Adkins R.M., Gibson S., Li W.H.;
RT   "Episodic Evolution of Growth Hormone in Primates and Emergence of the
RT   Species Specificity of Human Growth Hormone Receptor.";
RL   Mol. Biol. Evol. 18:[illegible] (2001).
DR   EMBL; AF339[illegible]; AAK[illegible];
DR   InterPro; IPR001050; Somatotropin.
DR   Pfam; PF00103; hormone; 1.
DR   PROSITE; PS00264; SOMATOTROPIN_1; UNKNOWN_1.
FT   NON_TER      1      1
FT   NON_TER    177    177
SQ   SEQUENCE   177 AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             [illegible]%;  Score [illegible];  DB 6;  Length 177;
  Best Local Similarity   [illegible]%;  Pred. No. [illegible];
  Matches  [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

QY       2 FPTIPLSRLFDNAMLRAHRLHQLAFDTY[illegible]KYS 44
           |||||||||| |||||  ||||| ||||||  |||||||||||
Db       4 FPAMSLSRLFANGVLRAQHLAADTFKEFERTYIPEGQRYS 48


RESULT 15
Q9Y5M6
ID   Q9Y5M6       PRELIMINARY;      PRT;   204 AA.
AC   Q9Y5M6;
DT   01-OCT-2000 (TrEMBLrel. 16, Created)
DT   01-OCT-2000 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Growth hormone (Fragment).
OS   Tursiops aduncus (Indian).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Odontoceti; Delphinidae; Tursiops.
OX   NCBI_TaxID=9739;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22065[illegible]; PubMed=[illegible];
RA   Liu J.Y., Makova K.D., Adkins R.M., Gibson S., Li W.H.;
RT   "Episodic Evolution of Growth Hormone in Primates and Emergence of the
RT   Species Specificity of Human Growth Hormone Receptor.";
RL   Mol. Biol. Evol. 18:[illegible] (2001).
DR   EMBL; AF339[illegible]; AAK[illegible];
DR   InterPro; IPR001050; Somatotropin.
DR   Pfam; PF00103; hormone; 1.
DR   PROSITE; PS00264; SOMATOTROPIN_1; UNKNOWN_1.
FT   NON_TER      1      1
FT   NON_TER    177    177
SQ   SEQUENCE   177 AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             [illegible]%;  Score [illegible];  DB 6;  Length 177;
  Best Local Similarity   [illegible]%;  Pred. No. [illegible];
  Matches  [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

QY       2 FPTIPLSRLFDNAMLRAHRLHQLAFDTY[illegible]KYS 44
           |||||||||| |||||  ||||| ||||||  |||||||||||
Db       8 FPSVPLSSLFANAVLRAQHLHQLAADTYKEFERTYIPEGQRYS 52


Search completed: February [illegible], 2003, [illegible]:[illegible]
Job time : [illegible] secs
```

Database: Published_Applications_AA:*

SUMMARIES

RESULT 1

RESULT 1
US-09-423-100-1
; Sequence 1, Application US/09423100
; Patent No. US6294672
; GENERAL INFORMATION:
; APPLICANT: Gan, Zhong Hua
; TITLE OF INVENTION: Chimeric Protein Containing an
; Intramolecular Chaperone Like Sequence
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/423,100
; FILING DATE: 02-Jan-2002
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/474,100
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/424,100
; FILING DATE: 21-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Frank J
; REGISTRATION NUMBER: 44,964
; REFERENCE/DOCKET NUMBER: 020676-000100US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; TELEX: 454508
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 92 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-423-100-1

Query Match 100.0%; Score 256; DB 0; Length 92;
Best Local Similarity 100.0%; Pred. No. 7.2e-28;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEFCIPLSKIFNAMLAAHCHLDAFLQYKEEAVIFKLKYSFLQNE 46
MEFCIPLSKIFNAMLAAHCHLDAFLQYKEEAVIFKLKYSFLQNE
Db 1 MEFCIPLSKIFNAMLAAHCHLDAFLQYKEEAVIFKLKYSFLQNE 46

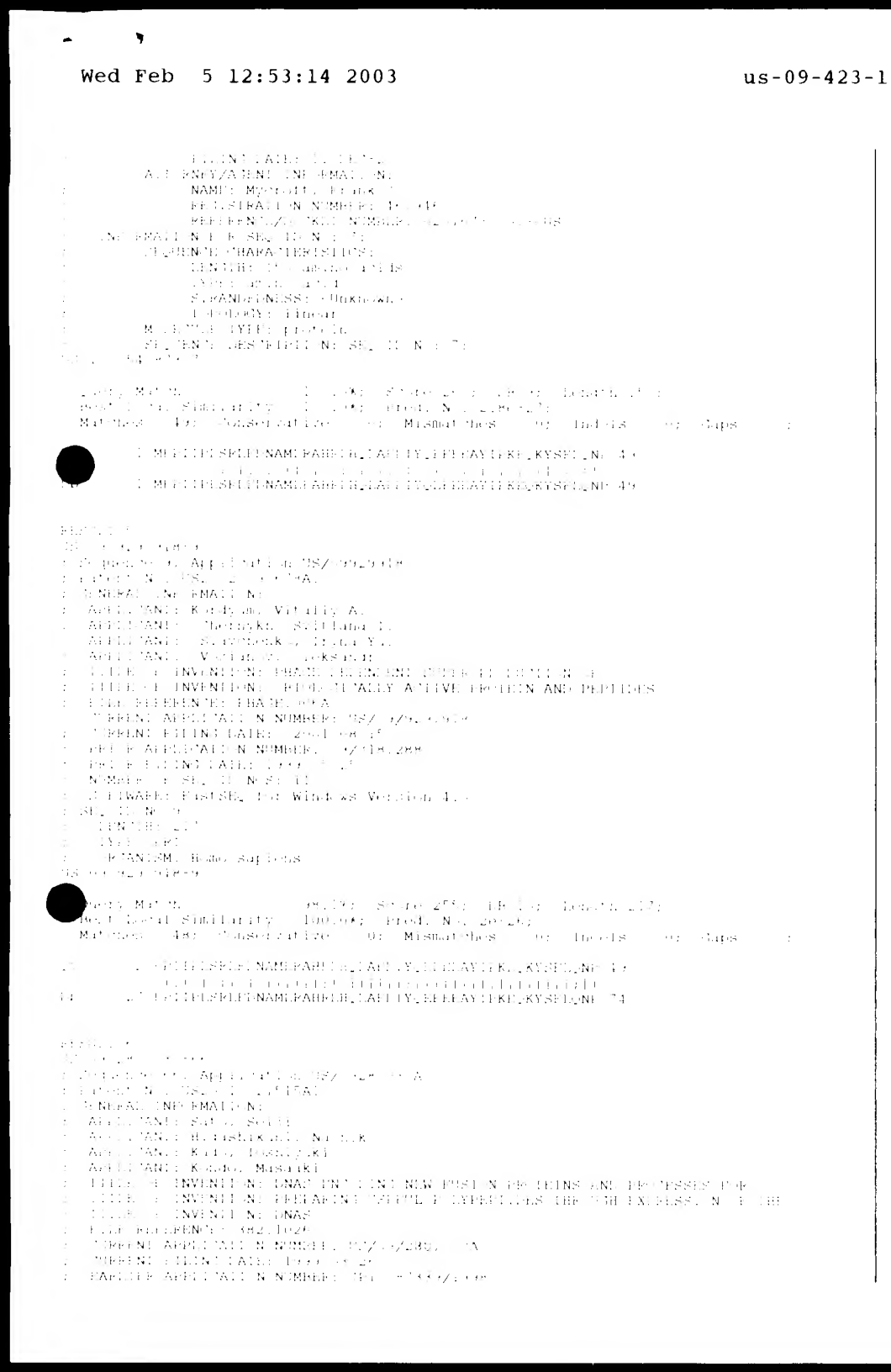
RESULT 2
US-09-423-100-2
; Sequence 2, Application US/09423100
; Patent No. US6294672
; GENERAL INFORMATION:
; APPLICANT: Gan, Zhong Hua
; TITLE OF INVENTION: Chimeric Protein Containing an
; Intramolecular Chaperone Like Sequence
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/423,100
; FILING DATE: 02-Jan-2002
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/474,100
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/424,100
; FILING DATE: 21-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Frank J
; REGISTRATION NUMBER: 44,964
; REFERENCE/DOCKET NUMBER: 020676-000100US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; TELEX: 454508
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 92 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-423-100-2

Query Match 100.0%; Score 256; DB 0; Length 92;
Best Local Similarity 100.0%; Pred. No. 7.2e-28;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEFCIPLSKIFNAMLAAHCHLDAFLQYKEEAVIFKLKYSFLQNE 46
MEFCIPLSKIFNAMLAAHCHLDAFLQYKEEAVIFKLKYSFLQNE
Db 1 MEFCIPLSKIFNAMLAAHCHLDAFLQYKEEAVIFKLKYSFLQNE 46

RESULT 3
US-09-423-100-3
; Sequence 3, Application US/09423100
; Patent No. US6294672
; GENERAL INFORMATION:
; APPLICANT: Gan, Zhong Hua
; TITLE OF INVENTION: Chimeric Protein Containing an
; Intramolecular Chaperone Like Sequence
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/423,100
; FILING DATE: 02-Jan-2002
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/474,100
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/424,100
; FILING DATE: 21-APR-1995



```
  Best Local Similarity   97.9%;  Pred. No. 1.7e-25;
  Matches   47;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        2 EFFIPLSKEFNAMLRARRIR,LAFCTY,EFRAYIFKEL,KYSFIL,NP 47
            |||||||||||||||||||||||||||||||||||||||||||||||
Db        2 EFFIPLSKEFNAMLRARRIR,LAFCTY,EFRAYIFKEL,KYSFIL,NP 47


RESULT 4
US-09-876-478-4
; Sequence 4, Application US/09876478
; Patent No. US20020168688A1
; GENERAL INFORMATION:
;  APPLICANT: PEI, DAVID N.
;  APPLICANT:  PFEIFER, ANNIE M.
;  APPLICANT:  SELGBY, JOHN
;  APPLICANT:  MCLEAN, DAVID S.
;  TITLE OF INVENTION: METHOD FOR DETECTING GROWTH HORMONE VARIATIONS IN
;  TITLE OF INVENTION: HUMANS, THE VARIANTS AND THEIR USES
;  FILE REFERENCE: WM79
;  CURRENT APPLICATION NUMBER: US/09/876,478
;  CURRENT FILING DATE: 2001-06-04
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 227
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-876-478-4

  Query Match             94.1%;  Score 342;  DB 10;  Length 217;
  Best Local Similarity   95.8%;  Pred. No. 3.2e-24;
  Matches   46;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        2 EFFIPLSKEFNAMLRARRIR,LAFCTY,EFRAYIFKEL,KYSFIL,NP 47
            | |||||||||||||||||||||||||||||||||||||||||||||
Db        2 EFFIPLSKEFNASLRARRIR,LAFCTY,EFRAYIFKEL,KYSFIL,NP 47


RESULT 5
US-09-876-478-5
; Sequence 5, Application US/09876478
; Patent No. US20020168688A1
; GENERAL INFORMATION:
;  APPLICANT: Pei, David
;  APPLICANT: Pfeifer, Annie M.
;  APPLICANT: Selgby, John
;  APPLICANT: McLean, David S.
;  TITLE OF INVENTION: METHOD FOR DETECTING GROWTH HORMONE VARIATIONS IN
;  TITLE OF INVENTION: HUMANS, THE VARIANTS AND THEIR USES
;  NUMBER OF SEQUENCES: 6
;  CORRESPONDENCE ADDRESS:
;  ADDRESSEE: INVITE PHARMACEUTICALS, INC.
;  STREET: 4721 BARRON DRIVE
;  CITY: PALO ALTO
;  STATE: CALIFORNIA
;  COUNTRY: USA
;  ZIP: 94304
;  COMPUTER READABLE FORM:
;  MEDIUM TYPE: Floppy disk
;  COMPUTER: IBM PC compatible
;  OPERATING SYSTEM: PC-DOS/MS-DOS
;  SOFTWARE: WordPerfect 5.1 for Windows 2.0/MS-DOS 5.0
;  CURRENT APPLICATION DATA:
;  APPLICATION NUMBER: US/09/876,478
;  FILING DATE: 07-Mar-2001
;  CLASSIFICATION: Unknown
;  PRIOR APPLICATION DATA:
;  APPLICATION NUMBER: US/60/187,852
;  FILING DATE: 08-Mar-2000
;  ATTORNEY/AGENT INFORMATION:
;  NAME: FERRONE, MICHAEL
;  REGISTRATION NUMBER: 37,132
;  REFERENCE/DOCKET NUMBER: PF-0545 US
;  TELECOMMUNICATION INFORMATION:
;  TELEPHONE: (650) 855-0555
;  TELEFAX: (650) 493-1209
;  INFORMATION FOR SEQ ID NO: 5:
;  SEQUENCE CHARACTERISTICS:
;  LENGTH: 217 amino acids
;  TYPE: amino acid
;  STRANDEDNESS: single
;  TOPOLOGY: linear
;  IMMEDIATE SOURCE:
;  LIBRARY: GenBank
;  CLONE: gi|4502027
;  SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-876-478-5

  Query Match             76.1%;  Score 277;  DB 10;  Length 217;
  Best Local Similarity   77.1%;  Pred. No. 4.7e-18;
  Matches   37;  Conservative    6;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        2 EFFIPLSKEFNAMLRARRIR,LAFCTY,EFRAYIFKEL,KYSFIL,NP 47
            | |||||| | |||||||||||||||||||||||||||||||||||||
Db        2 EFFIPLSKEFNASLRARRIRQLAFCTYQEFRAYIFKELDKYSFILENP 51


RESULT 6
US-09-876-478-14
; Sequence 14, Application US/09876478
; Patent No. US20020168688A1
; GENERAL INFORMATION:
;  APPLICANT: Pegoraro, John A.
;  TITLE OF INVENTION: Interferon Alpha Polypeptide Variant Gene
;  FILE REFERENCE: 7161
;  CURRENT APPLICATION NUMBER: US/09/876,478
;  CURRENT FILING DATE: 2001-06-04
;  PRIOR APPLICATION NUMBER: 60/210,092
;  PRIOR FILING DATE: 2000-06-07
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: MS-DOS
; SEQ ID NO 14
;  LENGTH: 91
;  TYPE: PRT
;  ORGANISM: HOMO SAPIENS
;  FEATURE:
;  OTHER INFORMATION: N TERMINAL RESIDUE IS MET
US-09-876-478-14

  Query Match             54.0%;  Score 197;  DB 10;  Length 91;
  Best Local Similarity   97.9%;  Pred. No. 2.2e-12;
  Matches   47;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        2 EFFIPLSKEFNAMLRARRIR,LAFCTY 47
            | |||||||||||||||||||||||||||||
Db       47 EFFIPLSKEFNASLRARRIRQLAFCTY 91


RESULT 12
US-09-876-478-2
; Sequence 2, Application US/09876478
; Patent No. US20020168688A1
; GENERAL INFORMATION:
;  APPLICANT: Pegoraro, John A.
;  TITLE OF INVENTION: Interferon Alpha Polypeptide Variant Gene
;  FILE REFERENCE: 7161
;  CURRENT APPLICATION NUMBER: US/09/876,478
;  CURRENT FILING DATE: 2001-06-04
;  PRIOR APPLICATION NUMBER: 60/210,092
;  PRIOR FILING DATE: 2000-06-07
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: MS-DOS
; SEQ ID NO 2
;  LENGTH: 34
;  TYPE: PRT
;  ORGANISM: HOMO SAPIENS
;  FEATURE:
```

Title: US-09-423-100-2

Perfect score: 470

Sequence: . MTEIIISSEIFNAMLSAHB.....NCSIIPISIII.SKSIRV. 2

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 3041726 seqs, 1142506262 residues

Total number of hits satisfying chosen parameters: 3041726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

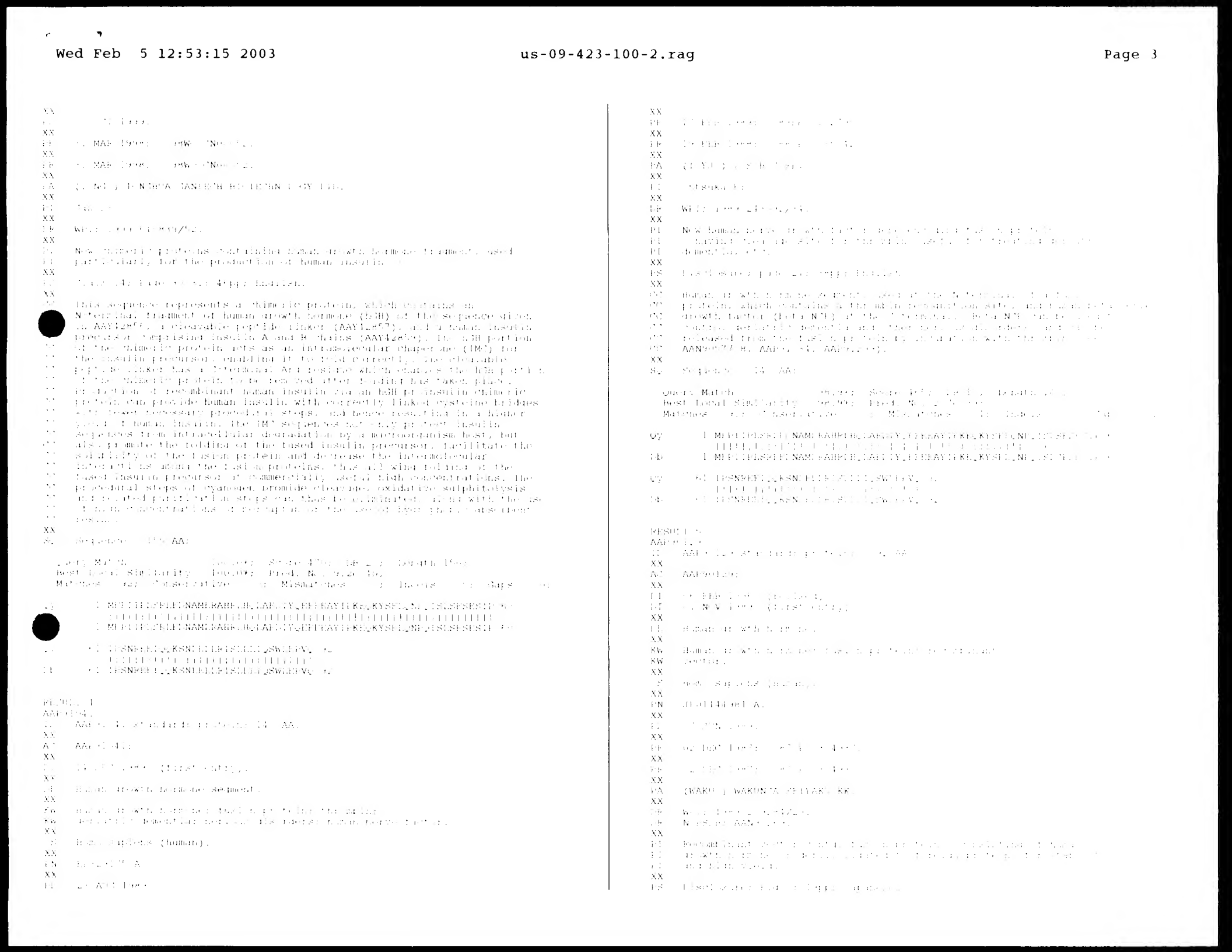
Database : A_Geneseq_101002:*

SUMMARIES

Result No. is the number of results produced by chance to have a score greater than or equal to the score of the result being printed, and is provided by analysis of the total score distribution.

ALIGNMENTS

```
PN   Claim 5; Page 28; 46pp; English.
XX
CC   The sequence represents an N-terminal fragment of human growth hormone
CC   (hGH) which is a component of a chimeric protein (AAY12[illegible]) which also
CC   includes a human insulin precursor (AAY12[illegible]). The hGH portion of the
CC   chimeric protein acts as an intramolecular chaperone (IMC) for the
CC   insulin precursor, enabling it to fold correctly. A cleavable peptide
CC   linker with a C-terminal Arg residue (AAY12[illegible]) enables the hGH portion
CC   of the chimeric protein to be removed after folding has taken place.
CC   Production of recombinant human insulin via an hGH-proinsulin chimeric
CC   protein can provide human insulin with correctly linked cysteine bridges
CC   with fewer necessary procedural steps, and hence resulting in a higher
CC   yield of human insulin. The IMC sequence not only protects insulin
CC   sequences from intracellular degradation by a microorganism host, but
CC   also promote the folding of the fused insulin precursor, facilitate the
CC   solubility of the fusion protein and decrease the intermolecular
CC   interactions among the fusion proteins. This allows voiding of the
CC   fused insulin precursor at commercially useful high concentrations. The
CC   procedural steps of cyanogen bromide cleavage, oxidative sulphitolysis
CC   and isolated purification steps can thus be eliminated, along with the
CC   use of high concentrations of mercaptan or the use of hydrophobic
CC   affinity resins.
XX
SQ   Sequence [illegible] AA;

  Query Match             100.0%;  Score 470;  DB 20;  Length 191;
  Best Local Similarity   100.0%;  Pred. No. [illegible];
  Matches  92;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy       1 MPTIFLFSMLFLNAMLAHGEQGTMTSPLFLFATIKSQKYSFTLNMRQKEEAYEKFLKIA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MPTIFLFSMLFLNAMLAHGEQGTMTSPLFLFATIKSQKYSFTLNMRQKEEAYEKFLKIA 60

Qy      61 IPSNREFFSLQKSNLELCHSLGLTLSWLEPVL 92
           |||||||||||||||||||||||||||||||
Db      61 IPSNREFFSLQKSNLELCHSLGLTLSWLEPVL 92


RESULT 2
AAW22665
ID   AAW22665 standard; protein; 191 AA.
XX
AC   AAW22665;
XX
DT   08-JUN-1999  (first entry)
XX
DE   Human anti-angiogenic protein 2/K hGH Met fusion.
XX
KW   Human; anti-angiogenic; prolactin; placental lactogen; hPL; angiogenesis;
KW   growth hormone; hGH; hGH V; capillary endothelial cell proliferation;
KW   pathological vascularisation; prognosis; treatment; angiogenic disease;
KW   cancer; tumour; metastasis; neovascular glaucoma; arteriosclerosis; arthritis;
KW   atherosclerotic plaques; corneal graft neovascularisation;
KW   wound healing; proliferative retinopathy; macular degeneration; trachoma;
KW   granulations; glaucoma; ocular; uveitis; fracture; osler-Weber syndrome;
KW   psoriasis; fibroplasia; scleroderma; Kaposi's sarcoma; vascular adhesions;
KW   chronic; [illegible]ma; reproductive disorder; contraception;
KW   dysfunction; pre-eclampsia; intrauterine growth retardation;
KW   placental dysfunction.
XX
OS   Homo sapiens.
XX
PN   WO9950302-A1.
XX
PD   07-OCT-1999.
XX
PF   01-MAY-1999;  99WO-US009641.
XX
PR   01-MAY-1998;  98US-0083984.
XX
PA   (GENE ) GENTECH INC.
XX
PI   Martin DA, Stroman L, Taylor R, Weiner RI;

XX
DR   WPI; [illegible]-[illegible]-Z.
DR   N-PSDB; AAX[illegible].
XX
PT   New and, anti-angiogenic peptides - [illegible] N-terminal fragment of
PT   human placental lactogen, human growth hormone, growth hormone
PT   variant or human prolactin.
XX
PS   Claim 5; Page 28; 46pp; English.
XX
CC   This invention describes novel human anti-angiogenic peptides derived
CC   from 10 to 150 consecutive amino acids selected from the N-terminal end
CC   of human placental lactogen (hPL), human growth hormone (hGH), an hGH
CC   hormone variant (hGH V), or human prolactin. Such peptides (i) inhibit
CC   capillary endothelial cell proliferation and organisation, (ii) inhibit
CC   angiogenesis in chick chorioallantoic membrane and (iii) bind to at
CC   least one specific receptor which does not bind an intact hPL, growth
CC   hGH, hPL, prolactin or hGH V. The invention also describes a method for
CC   diagnosing a probable abnormality of placental vascularisation during
CC   pregnancy. The peptides can be used for treating an angiogenic disease in
CC   a subject, for inhibiting tumour formation or growth in a patient, for
CC   modulating vascularisation in a patient's placenta, in pharmaceutical
CC   peptides can be used for prognostic or treatment of malignant and non-
CC   malignant tumours, arteriovenous malformation, arthritis such as rheumatoid
CC   arthritis, atherosclerotic plaques, corneal graft neovascularisation,
CC   delayed wound healing, proliferative retinopathy such as diabetic
CC   retinopathy, macular degeneration, transplants such as corneal
CC   implants, rapidly proliferating vascularised human cancers such as solid
CC   such as hypertrophic scars or keloid scars, vascular tumours, vascular
CC   tumour, occular neovascularisation, Osler Weber syndrome, pterygium
CC   pyogenic granuloma, retrolental fibroplasia, scleroderma, solid tumours,
CC   Kaposi's sarcoma, trachoma, vascular adhesions, chronic inflammation,
CC   leukaemia, and reproductive disorders such as infertility and dysfunction
CC   and pre-eclampsia. They can also be used as contraceptive agents. The
CC   circulating the peptides can be used in gene therapy. The [illegible] of
CC   abnormal levels of N-terminal fragments of hGH, hGH V, prolactin or hPL
CC   can be used in assays for impairment of vascular development associated
CC   with pre-eclampsia, intrauterine growth retardation, and placental
CC   dysfunction.
XX
SQ   Sequence [illegible] AA;

  Query Match             100.0%;  Score 470;  DB 20;  Length 191;
  Best Local Similarity   100.0%;  Pred. No. [illegible];
  Matches  92;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy       1 MPTIFLFSMLFLNAMLAHGEQGTMTSPLFLFATIKSQKYSFTLNMRQKEEAYEKFLKIA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MPTIFLFSMLFLNAMLAHGEQGTMTSPLFLFATIKSQKYSFTLNMRQKEEAYEKFLKIA 60

Qy      61 IPSNREFFSLQKSNLELCHSLGLTLSWLEPVL 92
           |||||||||||||||||||||||||||||||
Db      61 IPSNREFFSLQKSNLELCHSLGLTLSWLEPVL 92


RESULT 3
AAY12661
ID   AAY12661 standard; protein; [illegible] AA.
XX
AC   AAY12661;
XX
DT   15-JAN-2000  (first entry)
XX
DE   [illegible] [illegible] [illegible].
XX
KW   Insulin; precursor; growth hormone; chaperone; fusion protein;
KW   folding; conformation; chimeric protein; cleavable; recombinant;
KW   production; yield.
XX
OS   Homo sapiens.
XX
PN   WO9950302-A1.
XX
```



XX
PT The invention consists of a vector conta. a fusion protein which is
PT t ravel is, linarina, downstream of a promoter, hGH or a deri... (prot.
PT t med by subst'n. (Met 34 with Leu) and a foreign protein.
PT Stability of the vector in the host is greatly increased so the
PT pr tein yield is higher.
XX
SQ Sequence 192 AA;

Query Match 98.7%; Score 1[illegible]; DB 2; Length 192;
Best Local Similarity 98.9%; Pred. No. 3.9e-59;
Matches 91; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MEFIPTNLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSES 60
1 MEFIPTNLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSES 60

Qy 61 IPSNREETQQKSNLELLRISLLLIQSWLEPVQ 92
1 IPSNREETQQKSNLELLRISLLLIQSWLEPVQ 32

RESULT 2
AAW02254
ID AAW02254 standard; Protein; 192 AA.
XX
AC AAW02254;
XX
DT 14-NOV-1996 (first entry)
XX
DE Human anti-placental hormone peptide hGH Met-Ile-Leu-Val.
XX
KW Human; anti-placental, prolactin, placental lactogen, hCS, angiogenesis,
KW growth hormone; hGH; hGHV; capillary endothelial cell proliferation;
KW placental vascularisation; pregnancy; treatment; angiogenic disease;
KW tumour; inhibitor; malignant; antifibroma; arteriovenous malformation;
KW arthritis; atherosclerotic plaques; corneal graft neovascularisation;
KW w ound healing; proliferative retinopathy; macular degeneration; trachoma;
KW granulation; glaucoma; ocular; uveitis; fracture; osler Weber syndrome;
KW psoriasis; Hiv-plasia; scleroderma; Kaposi's sarcoma; vascular adhesion;
KW cancer; endometriosis; reproductive disorder; contraception;
KW ante therapy; pre-eclampsia; intrauterine growth retardation;
KW placental dysfunction.
XX
OS Homo sapiens.
XX
PN W 9[illegible] A1.
XX
PD 14-NOV-1996.
XX
PF 12-MAY-1995; 95WO-US06041.
XX
PR 13-MAY-1994; 94US-0242641.
XX
PA (REGE) UNIV CALIFORNIA.
XX
PI Martial JA, Struman I, Taylor SC, Weiner RI;
XX
DR WPI; 1996-4[illegible]/04.
DR N-PSDB; AAX02[illegible].
XX
PT New hGH, anti-placental peptides - comprise N-terminal fragments o
PT human placental lactogen, human gr. ... hormone, with ... substitutions.
PS Variant of human prolactin.
XX
CC Example 1; Page 6[illegible]; 8[illegible]pp; English.
XX
CC This invention describes a new human anti-placental specific peptides derived
CC from 2[illegible] to 15 consecutive amino acids selected from the N-terminal end
CC of human placental lactogen (hPL), human growth hormone (hGH), growth
CC hormone variant (hGHV), or human prolactin. Such peptides (i) inhibit
CC capillary endothelial cell proliferation and vascularisation (ii) initiate
CC angiogenesis in chick chorioallantoic membrane and (iii) binds to at
CC least one specific receptor which is not found on intact human cells
CC hGH, hPL, prolactin or hGHV. The invention also describes a method for
CC diagnosing a pregnancy abnormality, placental vascularisation, tumour
CC growth, The peptides are to used in treating an anti-fertility factor in
CC a subject, for inhibiting tumour formation or growth in a patient, and
CC modulating vascular wall in a patient's placenta, in particular the
CC peptides can be used for preventing treatment of malignant cancers,
CC angiofibroma, arteriovenous malformation, arthritis such as rheumatoid
CC arthritis, atherosclerotic plaques, corneal graft neovascularisation,
CC delayed wound healing, proliferative retinopathy, such as diabetic
CC retinopathy, macular degeneration, granulations such as those occurring
CC in haemophiliac joints, inappropriate vascularisation in wounded tissues
CC such as hepatic cirrhosis or Kaposi sarcoma, osteomyelitis, hemangioma,
CC psoriasis, uveitis, osler-weber syndrome, pyogenic granuloma, retrolental fibroplasia, scleroderma, solid tumours,
CC pyogenic glaucoma, retrolental fibroplasia, scleroderma, solid tumours,
CC Kaposi's sarcoma, trachoma, vascular adhesions, chronic inflammation,
CC leukaemia, and reproductive disorders such as follicular and luteal cysts
CC and choriocarcinoma. They can also be used as contraceptive agents. DNA
CC encoding the peptides can be used in gene therapy. The measurement of
CC abnormal levels of N-terminal fragments of hGH, hGHV, prolactin or hPL
CC can be used in assays for impairment of vascular development associated
CC with pre-eclampsia, inhibited placental growth retardation, and placental
CC dysfunction.
XX
SQ Sequence 192 AA;

Query Match 98.9%; Score 1[illegible]; DB 2; Length 192;
Best Local Similarity 98.9%; Pred. No. 3.9e-59;
Matches 91; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MEFIPTNLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSES 60
1 MEFIPTNLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSES 60

Qy 61 IPSNREETQQKSNLELLRISLLLIQSWLEIV 92
1 IPSNREETQQKSNLELLRISLLLIQSWLEIV 32

RESULT 3
AAE01[illegible]
ID AAE01[illegible] standard; protein; 201 AA.
XX
AC AAE01[illegible];
XX
DT 14-DEC-1999 (first entry).
XX
DE Human hormone growth factor and human growth hormone fusion protein.
XX
KW diagnostic; growth factor; fusion protein; therapy;
KW deficient dementia; nervous disorder; human growth hormone.
XX
OS Homo sapiens (human).
XX
FH Key Location/Qualifiers
FT Region 1..[illegible]
FT Region 141..154
FT Region 141..201
XX
PN EP[illegible]A.
XX
PD 2[illegible]-AUG-1995.
XX
PF 17-FEB-1989; 89EP-0102[illegible].
XX
PR 19-FEB-1988; 88JP-0037[illegible].
XX
PA (TOYO) TOYO JOZO KK.
XX
PI Ohtsuka E;
XX
DR WPI; 1989-2[illegible]/34.
XX

```
XX
KW   B-cell stimulatory factor-2; interleukin-6.
XX
OS   Homo sapiens.
XX
RN   [1] J Immunol 141:846-850,1988.
XX
RL   1-AUG-1988.
XX
DR   GENBANK; M23452.
XX
DR   GENBANK; M23452.
XX
RA   (TAYL ) TaSH Code.
XX
RT   Expression of human B cell differentiation factor 1 from specified DNA
RT   sequence segment, by recombinant DNA technique, gives protein or
RT   specified amino acid sequence.
XX
RL   Electronic Page 9: 17pp: Japanese.
XX
CC   The sequence encoding this protein can be fused with DNA encoding B cell
CC   differentiation factor (IL-6) and ligated into an expression vector for
CC   preparation of a fusion protein.
CC   See also AAB47411.
XX
SQ   Sequence  184 AA;

  Query Match             98.3%;  Score 462;  DB 11;  Length 184;
  Best Local Similarity   97.9%;  Pred. No. 2.2e-59;
  Matches  184;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 MFPGLTSPSLPTNAMEAAEQALSIALKTYDMTDEAYFKRQLKVSPLRMFSLQVASRDQK 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MFPGLTSPSLPTNAMEAAEQALSIALKTYDMTDEAYFKRQLKVSPLRMFSLQVASRDQK 60

Qy      61 TFSNMFELQVKSNLEELKSVLNLSKSVLHTVRPVL 96
             ||||||||||||||||||||||||||||||||||||
Db      61 TFSNMFELQVKSNLEELKSVLNLSKSVLHTVRPVL 96


RESULT 12
AAP81226
ID   AAP81226 standard; protein; 166 AA.
XX
AC   AAP81226;
XX
DT   20-NOV-2001  (first entry)
XX
DE   Sequence of protein with growth activity like a fragment.
XX
KW   Growth hormone.
XX
OS   Synthetic.
XX
DN   JP2001327286-A.
XX
PD   11-DEC-2001.
XX
PF   20-APR-2001;  2001JP-0121727.
XX
PR   20-APR-2000;  2000JP-0119203.
XX
PA   (TOYO ) TOYO SODA MFG KK.
XX
PI   Hirayoshi A;
DR   WPI; 2002-126815/17.
DR   N-PSDB; AAN81605.
XX
PT   Polypeptide with growth activity, useful
PT   in culturing bacterium transformed by plasmid vector, gene
PT   segment with specified DNA sequence
XX
PS   Claim 2(i); Page 5670; 9pp; Japanese.
XX
CC   The polypeptide (AA81225) with growth activity like a fragment of the GH
CC   (AAN81605) encoding it are examined. A Met residue may be added to
CC   the N terminal. The polypeptide acts in the bone structure and tissue
CC   including humans. It promotes bone growth with the polypeptide has high
CC   production rate and is easily extracted from bacterial cultures and
CC   utilized for use as a bone growth promoter.
XX
SQ   Sequence  166 AA;

  Query Match             [illegible]  Score [illegible];  DB 5;  Length [illegible];
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches  [illegible];  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MFPGLTSPSLPTNAMEAAEQALSIALKTYDMTDEAYFKRQLKVSPLRMFSLQVASRDQK 60

XX   Gene segment of human nerve growth factor - used in produ
XX   NGF production recombinant isolate, skin strain.
XX
PS   Claim 42; Page 482; 71pp; Japanese.
XX
CC   The protein is a direct transcript of the upstream 5'UTR gene
CC   promoter operator blocking its attenuation sequence and human
CC   beta NGF sequence. The gene fragment may be efficiently expressed in
CC   transformant host cell expression system.
CC   See also AAN60816-7.
XX
SQ   Sequence  242 AA;

  Query Match             [illegible]  Score [illegible];  DB 5;  Length [illegible];
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches  [illegible];  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MFPGLTSPSLPTNAMEAAEQALSIALKTYDMTDEAYFKRQLKVSPLRMFSLQVASRDQK 60

Db      61 TFSNMFELQVKSNLEELKSVLNLSKSVLHTVRPVL 96


RESULT 13
AAP81228
ID   AAP81228 standard; protein; 166 AA.
XX
AC   AAP81228;
XX
DT   20-NOV-2001  (first entry)
XX
DE   Sequence of protein with growth activity like a fragment.
XX
KW   Growth hormone.
XX
OS   Synthetic.
XX
DN   JP2001327286-A.
XX
PD   11-DEC-2001.
XX
PF   20-APR-2001;  2001JP-0121727.
XX
PR   20-APR-2000;  2000JP-0119203.
XX
PA   (TOYO ) TOYO SODA MFG KK.
XX
PI   Hirayoshi A;
DR   WPI; 2002-126815/17.
DR   N-PSDB; AAN81605.
XX
PT   Polypeptide with growth activity, useful
PT   in culturing bacterium transformed by plasmid vector, gene
PT   segment with specified DNA sequence
XX
PS   Claim 2(i); Page 5670; 9pp; Japanese.
XX
CC   The polypeptide (AA81227) with growth activity like a fragment of the GH
CC   (AAN81605) encoding it are examined. A Met residue may be added to
CC   the N terminal. The polypeptide acts in the bone structure and tissue
CC   including humans. It promotes bone growth with the polypeptide has high
CC   production rate and is easily extracted from bacterial cultures and
CC   utilized for use as a bone growth promoter.
XX
SQ   Sequence  188 AA;

  Query Match             [illegible]  Score [illegible];  DB 5;  Length [illegible];
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches  [illegible];  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 TFPGLTSPSLPTNAMEAAEQALSIALKTYDMTDEAYFKRQLKVSPLRMFSLQVASRDQK 61
```

PA (SUMI) SUMITOMO SEIYAKU KK.
XX
RN Wpi; 1999-284567/24.
XX
PT A human growth hormone mutant - with equivalent activity to natural
PT human growth hormone
XX
PS Example 1; Page 5-6; 10pp; Japanese.
XX
CC The present invention describes a human growth hormone mutant in which
CC the 134th Arg and the 135th Thr are replaced respectively by Asp and Pro
CC in the 1st to the 191st amino acid sequence of natural type human 22 kDa
CC growth hormone (hGH) and which has a resistance against decomposition by
CC thrombin. The present sequence represents the natural hGH. Also
CC described are: (1) a hGH mutant in which the 134th Arg, the 135th Thr
CC and the 140th Lys are replaced respectively by Asp, Pro and Ala in the
CC amino acid sequence of natural type hGH and which has a resistance
CC against decomposition by thrombin and plasmin; and (2) a drug
CC preparation containing the above hGH mutant as the active component.
CC The mutant hGH shows an activity approximately equivalent to that of
CC natural type hGH and shows a high stability in blood and body fluid.
XX
SQ Sequence 191 AA;

Query Match 97.9%; Score 460; DB 20; Length 191;
Best Local Similarity 98.9%; Pred. No. 1.2e-38;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPT 60
||:|||||:|:|||||||||:|||||||||||||||||:|||||||:|:|||| ||||| |||||||
Db 1 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPT 60

Qy 61 PSNREETQQKSNLELLRISLLLIQSWLEPVQ 92
|||||||:|||||||||||||||||:|:|||||||
Db 61 PSNKEETQQKSNLELLRISLLLIQSWLEPVQ 91

Search completed: February 5, 2003, 12:44:26
Job time : 38.5914 secs

[illegible] protein search, using sw model

[illegible]

Title: US-09-423-100-2
[illegible] 170
[illegible] 1 MFLLIPCSHLFLNAMLRAHR..........NCLLRG[illegible]SWIFPV [illegible]

[illegible] BLOSUM62
[illegible]

[illegible]

[illegible]

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

[illegible] analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
[illegible]	[illegible]	[illegible]	[illegible]	1	STHU	somatotropin [illegible]
[illegible]	[illegible]	[illegible]	[illegible]	2	[illegible]	somatotropin [illegible]
[illegible]	[illegible]	[illegible]	[illegible]	1	STHUV	somatotropin [illegible]
[illegible]	[illegible]	[illegible]	[illegible]	1	STHUV2	somatotropin [illegible]
[illegible]	[illegible]	[illegible]	[illegible]	2	[illegible]	somatotropin [illegible]
[illegible]	[illegible]	[illegible]	[illegible]	2	[illegible]	chorionic somatomam
[illegible]	[illegible]	[illegible]	[illegible]	2	[illegible]	chorionic somatomam
[illegible]	[illegible]	[illegible]	[illegible]	2	[illegible]	chorionic somatomam
[illegible]	[illegible]	[illegible]	217	1	[illegible]	choriomammotropin
[illegible]	[illegible]	[illegible]	217	2	[illegible]	choriomammotropin
[illegible]	[illegible]	[illegible]	[illegible]	2	A26449	choriomammotropin
[illegible]	[illegible]	[illegible]	[illegible]	2	[illegible]	somatotropin [illegible]
[illegible]	[illegible]	[illegible]	[illegible]	2	[illegible]	somatotropin [illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	STMS	somatotropin [illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	somatotropin [illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	somatotropin [illegible]
[illegible]	[illegible]	[illegible]	[illegible]	2	[illegible]	somatotropin [illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	somatotropin [illegible]
[illegible]	[illegible]	[illegible]	[illegible]	1	[illegible]	somatotropin precu
[illegible]	[illegible]	[illegible]	[illegible]	2	[illegible]	somatotropin precu
[illegible]	[illegible]	[illegible]	[illegible]	2	[illegible]	somatotropin precu
[illegible]	[illegible]	[illegible]	[illegible]	2	A37792	somatotropin precu
[illegible]	[illegible]	[illegible]	[illegible]	1	A[illegible]	somatotropin [illegible]
[illegible]	[illegible]	[illegible]	[illegible]	2	[illegible]	somatotropin [illegible]
[illegible]	[illegible]	[illegible]	217	1	[illegible]	somatotropin precu
[illegible]	[illegible]	[illegible]	217	[illegible]	[illegible]	somatotropin precu
[illegible]	[illegible]	[illegible]	217	[illegible]	[illegible]	somatotropin precu
[illegible]	[illegible]	[illegible]	217	2	[illegible]	somatotropin [illegible]
[illegible]	[illegible]	[illegible]	[illegible]	2	[illegible]	somatotropin precu

[illegible]

ALIGNMENTS

RESULT 1
STHU
somatotropin [illegible]
N;Alternate names: [illegible]
N;Contains: growth hormone [illegible]
C;Species: Homo sapiens (man)
C;Date: 24 Apr 1984 #sequence_revision [illegible] #text_change [illegible]
C;Accession: [illegible]
R;DeNoto, F.M.; Moore, D.D.; Goodman, H.M.
Nucleic Acids Res. [illegible]
A;Title: Human growth hormone DNA sequence and mRNA structure: [illegible]
A;Reference number: [illegible]
A;Accession: [illegible]
A;Molecule type: DNA
A;Residues: [illegible]
A;Cross-references: [illegible]
A;Note: the 20K [illegible]
R;Chen, E.Y.; Liao, Y.C.; Smith, D.H.; Barrera-Saldana, H.A.; [illegible]
Genomics 4, 479-497, 1989
A;Title: The human growth hormone locus: nucleotide sequence, [illegible]
A;Reference number: [illegible]
A;Accession: [illegible]
A;Molecule type: DNA
A;Residues: [illegible]
A;Cross-references: [illegible]
R;Roskam, W.; [illegible]
Nucleic Acids Res. [illegible]
A;Title: [illegible]
A;Reference number: [illegible]
A;Accession: [illegible]
A;Molecule type: mRNA
A;Residues: [illegible]
A;Cross-references: [illegible]
A;Note: [illegible]
R;Martial, J.A.; Hallewell, R.A.; Baxter, [illegible]
Science [illegible]
A;Title: Human growth hormone: [illegible]
A;Reference number: [illegible]
A;Accession: [illegible]
A;Molecule type: mRNA
A;Residues: [illegible]
R;[illegible]
Arch. Biochem. Biophys. [illegible]
A;Title: Human pituitary growth hormone. [illegible]
A;Reference number: [illegible]
A;Contents: annotation
R;Li, C.H.; [illegible]
Arch. Biochem. Biophys. [illegible]
A;Title: Human pituitary growth hormone. [illegible]
A;Reference number: [illegible]
A;Accession: [illegible]
A;Molecule type: protein

A;Residues: 27-94;96-217 <LIC>
R;Niall, H.D.
Nature New Biol., 230, 90-91, 1971
A;Title: Revised primary structure for human growth hormone.
A;Reference number: A93793; MUID:71179755; PMID:5273618
A;Accession: A93793
A;Molecule type: protein
A;Residues: 27-51<NIA>
R;Niall, H.D.; Hogan, M.L.; Sauer, R.; Rosenblum, I.Y.; Greenwood, F.C.
Proc. Natl. Acad. Sci. U.S.A. 68, 866-869, 1971
A;Title: Sequences of pituitary and placental human growth hormones, porcine
A;Reference number: A93779; MUID:71127965; PMID:5279526
A;Accession: A93779
A;Molecule type: protein
A;Residues: 119-120;157-159 <NI2>
R;Niall, H.D.
In Prolactin and Carcinogenesis, Proc. Fourth Tenovus Workshop, Griffiths, K.,
A;Title: The chemistry of the human lactogenic hormones.
A;Reference number: A94427
A;Contents: annotation; somatotropin revi...
R;Rowley, L.A.; Dixon, J.S.; Li, C.H.
Int. J. Pept. Protein Res. 4, 281-287, 1972
A;Title: Sequence comparison of human pituitary growth hormone, human chorionic somatom
A;Reference number: A91761; MUID:73092029; PMID:4679454
A;Accession: A91764
A;Molecule type: protein
A;Residues: 27-217 <BEW>
R;Beattie, C.W.; Rosenfeld, L.F.; Lewis, L.J.
Biochem. Biophys. Res. Commun. 92, 511-516, 1980
A;Title: The 22,000-dalton variant of human growth hormone: location of the amino acid d
A;Reference number: A90217; MUID:80170196; PMID:7356479
A;Contents: somatotropin, 20K short variant
A;Accession: A90217
A;Molecule type: protein
A;Residues: 48-57;73-80;...-183
R;Chapman, G.E.; Rogers, K.M.; Brittain, T.; Bradshaw, R.A.; Potts, J.T., Jr.; Niall,
J. Biol. Chem. 256, 2395-2401, 1981
A;Title: The 20,000 molecular weight variant of human growth hormone: Preparation and
A;Reference number: A92411; MUID:81117461; PMID:7462247
A;Contents: somatotropin, 20K short variant
A;Accession: A92411
A;Molecule type: protein
A;Residues: 27-57;73-79 <CHA>
R;Singh, B.N.P.; Gaffney, B.K.; Lewis, L.D.; Lewis, U.J.
J. Protein Chem. 2, 425-436, 1983
A;Title: Human growth hormone peptide 1-43: isolation from pituitary glands.
A;Reference number: A81466
A;Accession: A81466
A;Molecule type: protein
A;Residues: 27-69<SIN>
A;Note: growth hormone 5K peptide has insulin potentiating activity; its physiological
R;Kousba, V.M.J.; Rao, I.D.; NG, B.
Biol. Chem. Hoppe-Seyler 371, 423-431, 1990
A;Title: Identification of the aspartimide structure in a previously reported peptide.
A;Reference number: S09605; MUID:90134745; PMID:2175679
A;Accession: S09605
A;Molecule type: protein
A;Residues: 27-44,'D',46-47 <BOB>
R;de Vos, A.M.; Ultsch, M.; Kossiakoff, A.A.
Science 255, 306-312, 1992
A;Title: Human growth hormone and extracellular domain of its receptor: crystal structur
A;Reference number: A41728; MUID:92196572; PMID:1549776
A;Contents: annotation; X-ray crystallography, 2.8 angstroms
A;Note: the structure of the complex with an rGH hormone receptor is described
R;Gray, G.L.; Baldridge, J.S.; McKeown, K.S.; Heynoker, H.L.; Chang, C.N.
Gene 39, 247-254, 1985
A;Title: Periplasmic production of correctly processed human growth hormone in Escherich
A;Reference number: 141126; MUID:86127293; PMID:3012261
A;Accession: I84549
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-26 <RES>
A;Cross-references: GB:M11568; NID:g178...; PID:AAA52654.1; PID:g17...

C;Comment: The gene for this hormone is transcribed only in somatotrophs [illegible]
C;Comment: About 90% of somatotropin is the 22K form.
C;Genetics:
A;Gene: GDB:GH1
A;Cross-references: GDB:119...; OMIM:17...
A;Map position: 17q23.1-17q24.3
A;Introns: 4/...; 57/...; 97/...; 155/...
C;Superfamily: prolactin
C;Keywords: alternative splicing; hormone; pituitary
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-217/Product: somatotropin [...] #status experimental <MAT>
F;27-48/Disulfid: growth hormone 5K peptide #status experimental <5KP>
F;27-57,73-217/Product: somatotropin 20K short form #status experimental <SHO>
F;79-191,...-217/Disulfid: [illegible] #status experimental

  Query Match 97.9%; Score 460; DB 1; Length 217;
  Best Local Similarity 98.6%; Pred. No. 1.6e-42;
  Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPTP 61
||||||||||||||||||||| |||||||||||||||||||||||||||||| ||||
Db 27 FPTIPLSGLFANAVLRAQHLHQLAADTFKEFERTYIPEGQRYSIQNTQVAFCFSETIPAP 86

Qy 62 PSNKEETQQKSNLELLRISLLLIQSWLEPVQ 92
|||||||||||||||||||||||||||||||
Db 87 PSNREETQQKSNLELLRISLLLIQSWLEPVQ 117

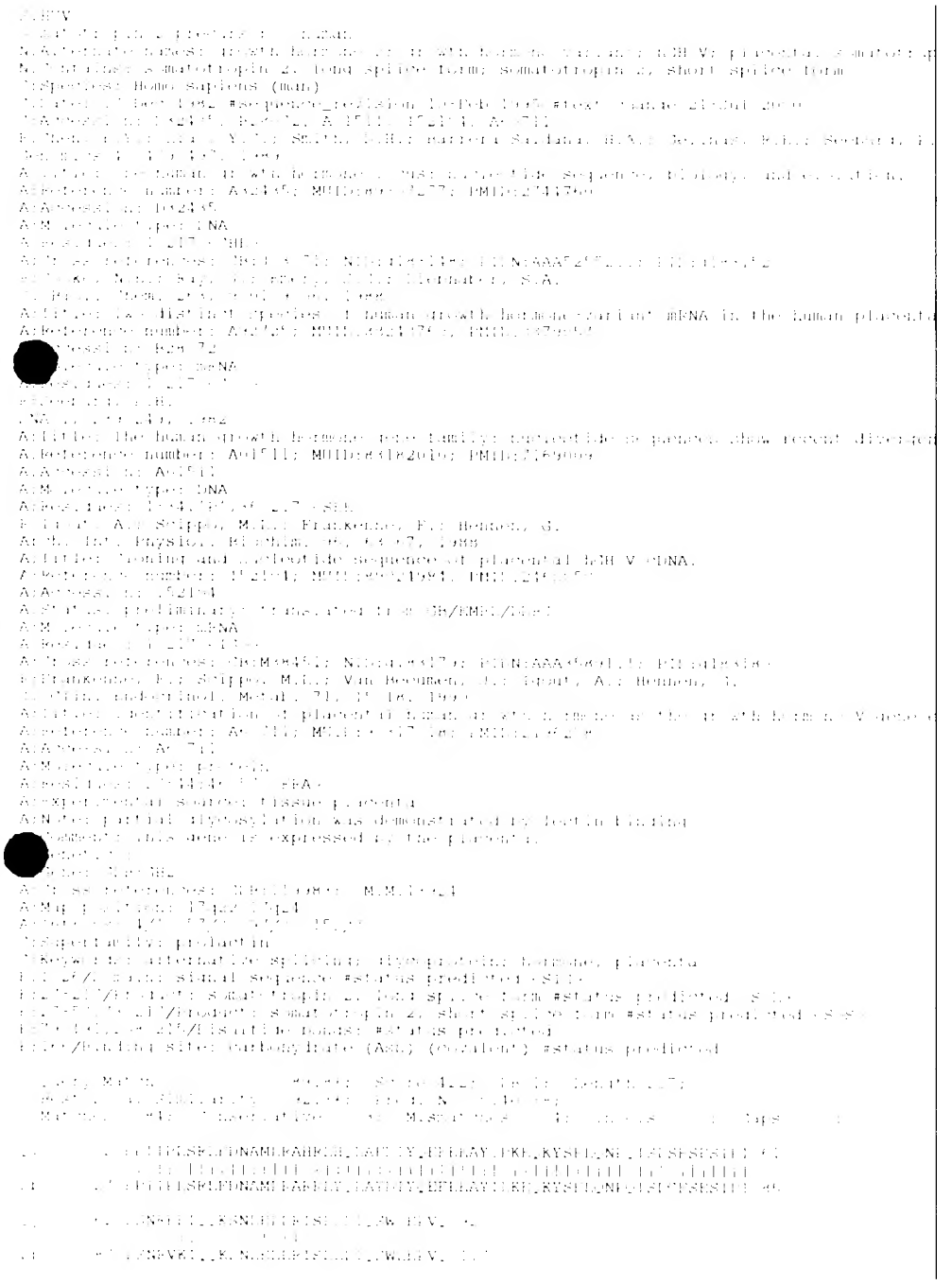
RESULT 2
167410
somatotropin - rhesus macaque
N;Alternate names: growth hormone
C;Species: Macaca mulatta (rhesus macaque)
C;Date: 31-May-1996 #sequence revision 31-May-1996 #text change 16-Jul-1999
C;Accession: I67410; A05094
R;Cubas, Jr.; Goodarzi, M.; Fiedler, J.H.; Fuller, P.D.
Endocrinology 133, 1744-1752, 1993
A;Title: Cloning of [illegible]/chorionic somatomammotropin related complemen
A;Reference number: I53967; MUID:94039824; PMID:8404617
A;Accession: I67410
A;Status: translated from GenBank/EMBL
A;Molecule type: mRNA
A;Residues: 1-217 <RES>
A;Cross-references: GB:L16556; NID:g...; PID:AAA36642.1; PID:g...
R;Li, C.H.; Chung, D.; Lahm, H.W.; Stein, S.
Arch. Biochem. Biophys. 245, 287-291, 1986
A;Title: The primary structure of monkey pituitary growth hormone.
A;Reference number: A05094; MUID:86129460; PMID:3080969
A;Accession: A05094
A;Molecule type: protein
A;Residues: 27-99,'Q',101-178,'D',180-217 <LIC>
A;Note: the monkey species is not identified in the reference
R;Raben, M.S.
Science 125, 883-884, 1957
A;Title: Preparation of growth hormone from pituitaries of man and monkey
A;Reference number: A41774
A;Contents: annotation; identification of source organism
C;Superfamily: prolactin

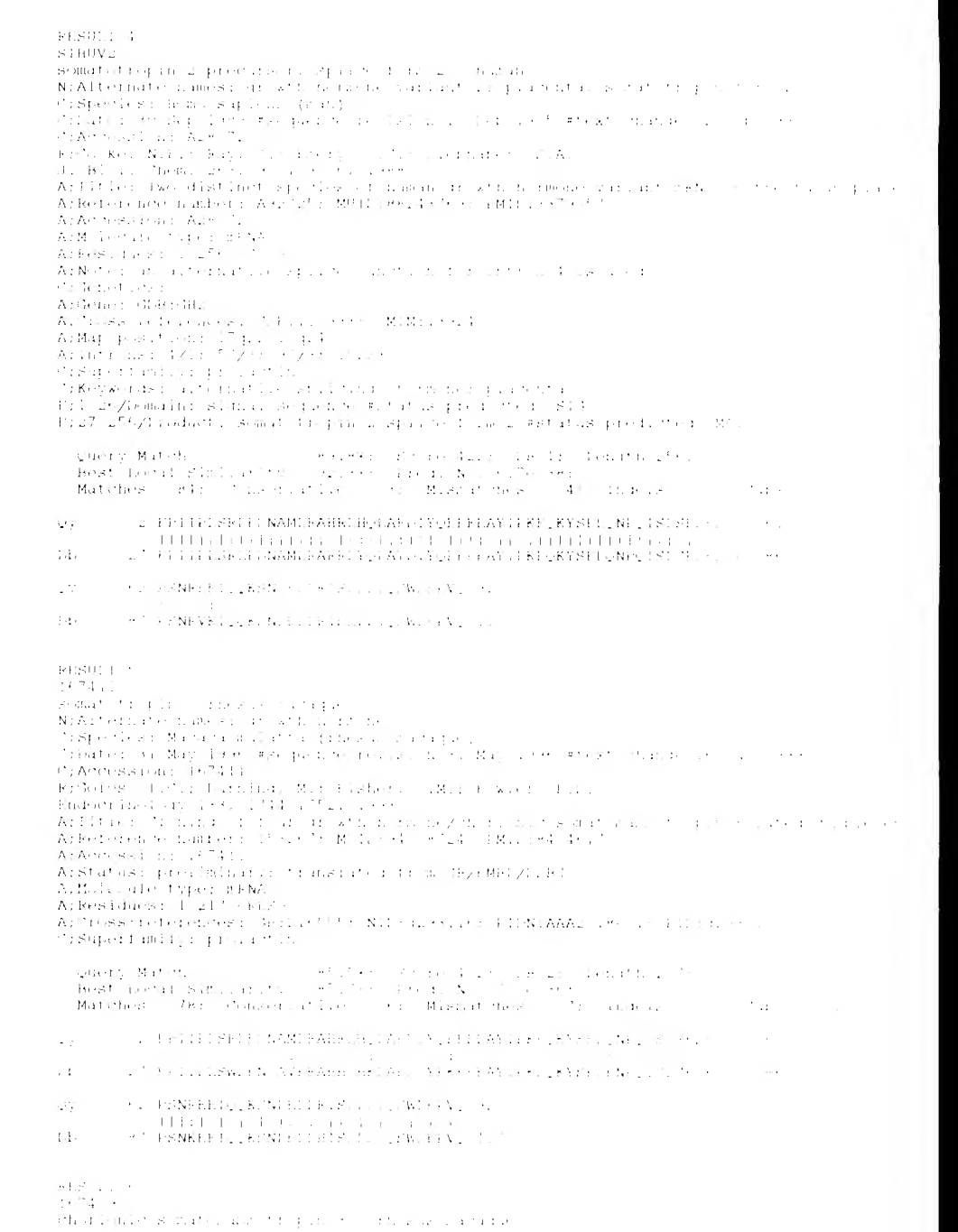
  Query Match 97.9%; Score 460; DB 2; Length 217;
  Best Local Similarity 98.6%; Pred. No. 1.6e-42;
  Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

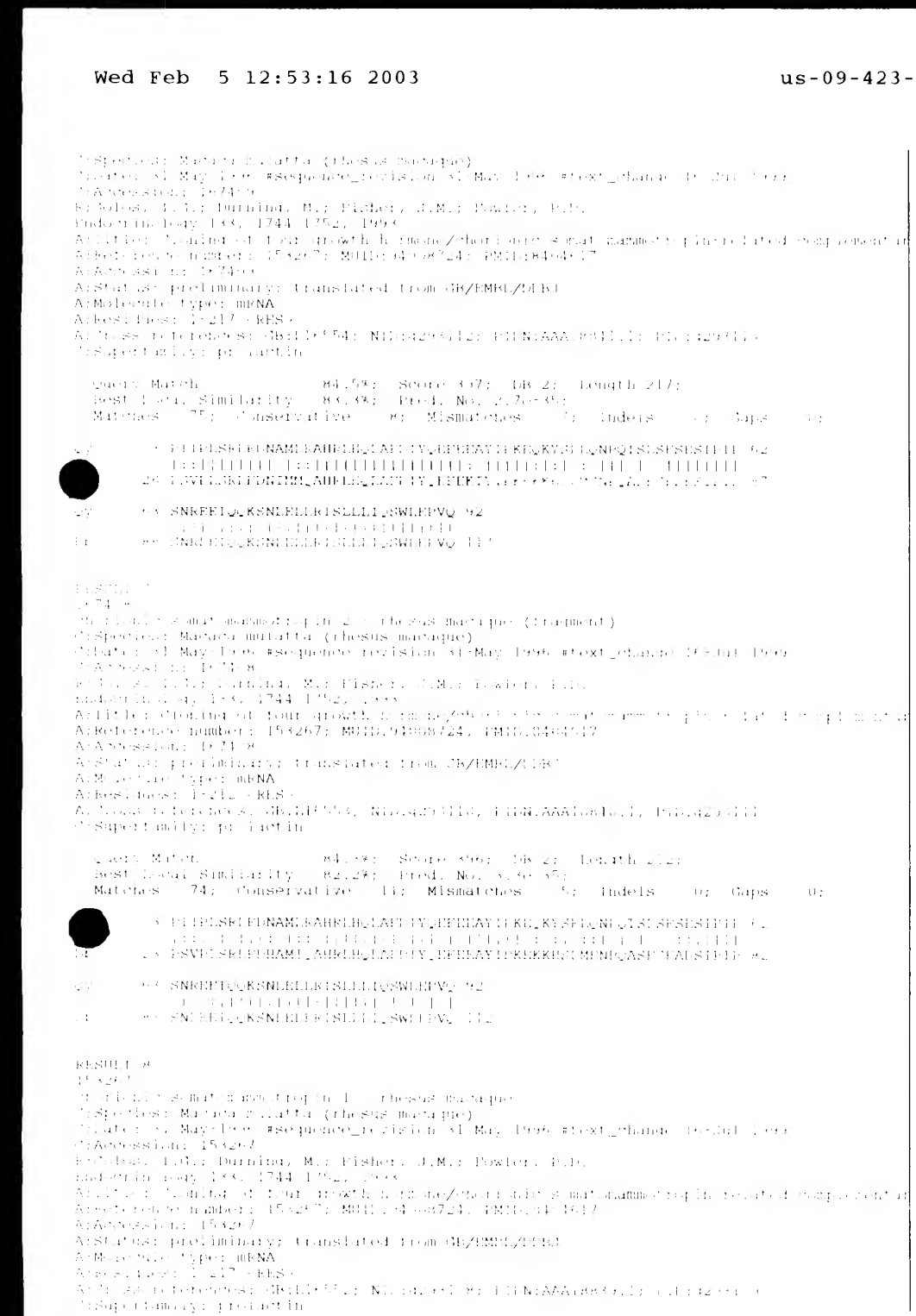
Qy 2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPTP 61
||||||||||||||||||||| |||||||||||||||||||||||||||||| ||||
Db 27 FPTIPLSGLFANAVLRAQHLHQLAADTFKEFERTYIPEGQRYSIQNTQVAFCFSETIPAP 86

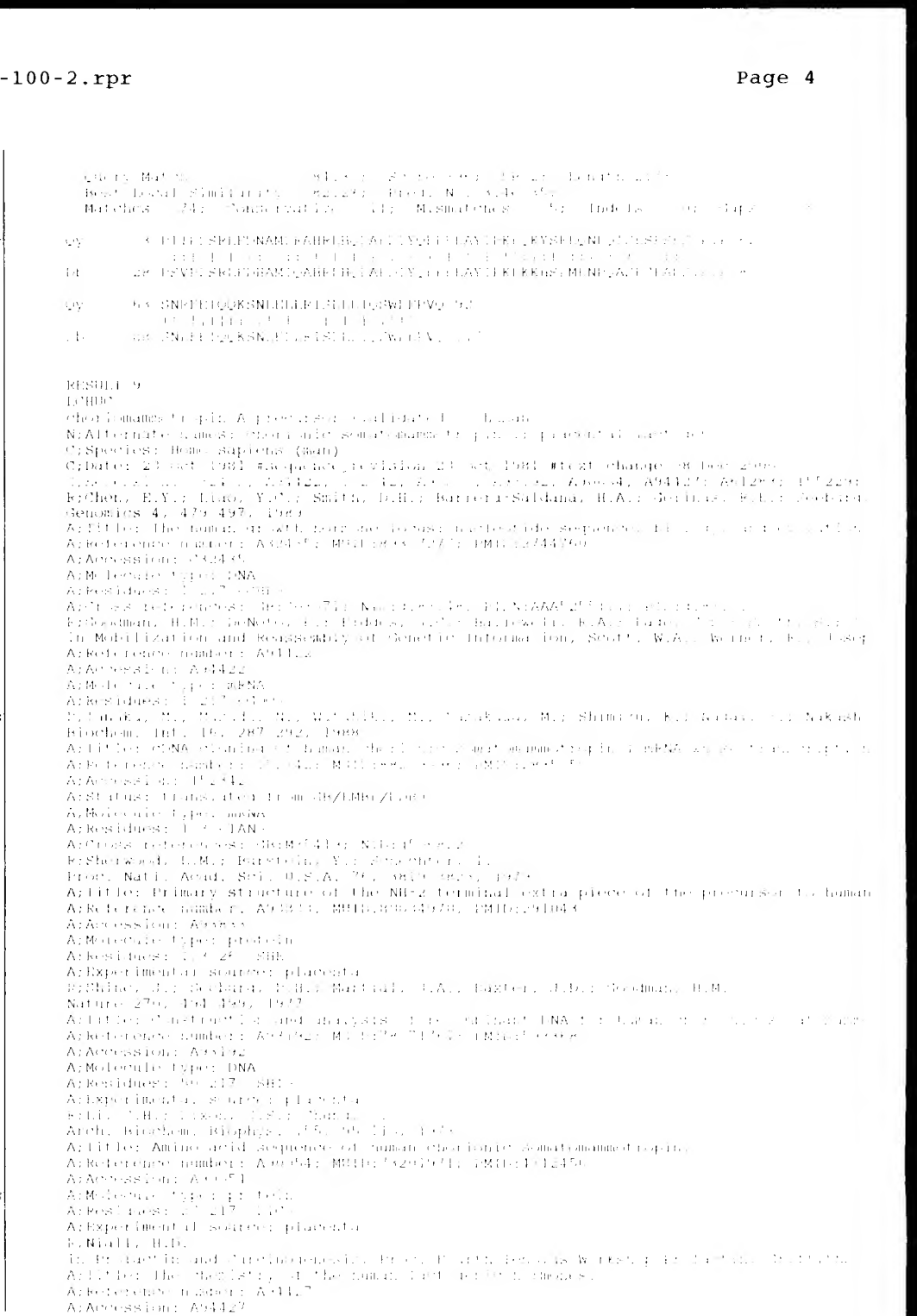
Qy 62 PSNKEETQQKSNLELLRISLLLIQSWLEPVQ 92
|||||||||||||||||||||||||||||||
Db 87 PSNREETQQKSNLELLRISLLLIQSWLEPVQ 117

RESULT









```
A;Accession: B491[illegible]
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-216 <SOU>
A;Cross-references: GB:S662[illegible], NID:g2[illegible], PIDN:AAB2[illegible].1, PID:g2[illegible]56
A;Note: sequence extracted from NCBI backbone (NCBIN:[illegible], NCBIP:[illegible])
C;Superfamily: prolactin

  Query Match             66.1%;  Score 310.5;  DB 2;  Length 216;
  Best Local Similarity   67.0%;  Pred. No. 5.4e-26;
  Matches   61;  Conservative   1[illegible];  Mismatches   1[illegible];  Indels    1;  Gaps    1;

Qy        2 FPTIPLSRLF[illegible]NAML[illegible]SFSESIPT 61
Db       27 FPAM[illegible]A 8[illegible]

Qy       62 PSNREEIQQKSNLELLRISLLLIQSWLEPVQ 92
Db       8[illegible] P[illegible]KEEA[illegible]MELLRPSLLLIQSWL[illegible]PVQ 11[illegible]


RESULT 1[illegible]
FN0140
somatotropin - sei whale
N;Alternate names: growth hormone
C;Species: Balaenoptera borealis (sei whale)
C;Date: 07-May-1999 #sequence_revision 07-May-1999 #text_change 07-May-1999
C;Accession: FN0140
R;Yudaev, N.A.; [illegible]
Biokhimiia 47, 1059-1069, 1982
A;Title: Amino acid sequence of seiwhale somatotropin.
A;Reference number: FN014[illegible]; MUID:8[illegible]; PMID:7[illegible]
A;Accession: FN0140
A;Molecule type: protein
A;Residues: 1-190 <YUD>
A;Note: article in Russian with English abstract
C;Superfamily: prolactin
C;Keywords: growth factor; hormone
F;52-16[illegible],18[illegible]-188/Disulfide bonds: #status predicted

  Query Match             65.4%;  Score 307.5;  DB 2;  Length 190;
  Best Local Similarity   67.0%;  Pred. No. 9.8e-26;
  Matches   61;  Conservative   14;  Mismatches   15;  Indels    1;  Gaps    1;

Qy        2 FPTIPLSRLF[illegible]NAML[illegible] 61
Db        1 FPAM[illegible]ANAVLRAQHL[illegible]VTL[illegible] 5[illegible]

Qy       62 PSN[illegible] 92
Db       6[illegible] PANK[illegible]RAQ[illegible]RSDVELLRPSLLLIQSWLGPVQ 90


RESULT 14
S1MS
somatotropin precursor - mouse
N;Alternate names: growth hormone
C;Species: Mus musculus (house mouse)
C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 28-May-1999
C;Accession: B2[illegible]
R;Linzer, D.I.H.; Talamantes, F.
J. Biol. Chem. 260, 9574-9579, 1985
A;Title: Nucleotide sequence of mouse prolactin and growth hormone mRNAs and expression
A;Reference number: A[illegible]; MUID:[illegible]; PMID:[illegible]
A;Accession: B23911
A;Molecule type: mRNA
A;Residues: 1-216 <LIN>
A;Cross-references: GB:X0[illegible], [illegible], NID:g[illegible], PIDN:[illegible].1, PID:g[illegible]
C;Superfamily: prolactin
C;Keywords: anterior pituitary; growth factor; hormone
F;1-2[illegible]/Domain: signal sequence #status predicted <SIG>
F;27-21[illegible]/Product: somatotropin #status predicted <SIN>
F;78-18[illegible],20[illegible]-214/Disulfide bonds: #status predicted

  Query Match             64.8%;  Score 304.5;  DB 1;  Length 21[illegible];
  Best Local Similarity   64.8%;  Pred. No. 2.4e-25;
  Matches   59;  Conservative   1[illegible];  Mismatches   17;  Indels    1;  Gaps    1;

Qy        2 [illegible]NAML[illegible]NQ[illegible]SLSFSESIPT 61
Db       27 FPAM[illegible]SAAFCFSETIPA 8[illegible]

Qy       62 PSNREEIQQKSNLELLRISLLLIQSWLEPVQ 92
Db       86 PTGKEEAQQR[illegible]MELLRFSLLLIQSWLGPVQ 116


RESULT 15
STHO
somatotropin - horse
N;Alternate names: growth hormone
C;Species: Equus caballus (domestic horse)
C;Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change [illegible]
C;Accession: A91772; A9[illegible]; A9[illegible]; A9[illegible]; A01514
R;Zakin, M.M.; Poskus, E.; Langton, A.A.; Ferrara, P.; Santome, J.A.; Dellacha, J.M.;
Int. J. Pept. Protein Res. 8, 435-444, 1976
A;Title: Primary structure of equine growth hormone.
A;Reference number: A91772; MUID:77005410; PMID:965151
A;Accession: A91772
A;Molecule type: protein
A;Residues: 1-190 <ZAK>
R;Zakin, M.M.; Poskus, E.; Dellacha, J.M.; Paladini, A.C.; Santome, J.A.
FEBS Lett. 34, 353-355, 1973
A;Title: The amino acid sequence of equine growth hormone.
A;Reference number: A91395; MUID:74[illegible]; PMID:4747849
A;Accession: A91395
A;Molecule type: protein
A;Residues: 1-190 <ZA2>
R;Zakin, M.M.; Poskus, E.; Dellacha, J.M.; Paladini, A.C.; Santome, J.A.
FEBS Lett. 25, 77-82, 1972
A;Title: Amino acid sequence around the cystine residues in equine growth hormone.
A;Reference number: A91383
A;Accession: A91383
A;Molecule type: protein
A;Residues: 42-69;157-190 <ZA3>
R;Oliver, L.; Hartree, A.S.
Biochem. J. 109, 19-24, 1968
A;Title: [illegible]
A;Reference number: A90240; MUID:68368058; PMID:4876100
A;Accession: A90240
A;Molecule type: protein
A;Residues: 176-190 <OLI>
C;Superfamily: prolactin
C;Keywords: hormone; pituitary
F;52-163,180-188/Disulfide bonds: #status experimental

  Query Match             64.4%;  Score 302.5;  DB 1;  Length 190;
  Best Local Similarity   64.8%;  Pred. No. 3.4e-25;
  Matches   59;  Conservative   14;  Mismatches   17;  Indels    1;  Gaps    1;

Qy        2 [illegible]NAML[illegible]NQ[illegible]SLSFSESIPT 61
Db        1 FPAM[illegible]AAFCFSETIPA 59

Qy       62 PSNREEIQQKSNLELLRISLLLIQSWLEPVQ 92
Db       6[illegible] P[illegible]KDEA[illegible]MELLRFSLLLI[illegible]SWL[illegible]PVQ 90


Search completed: February [illegible], 2003, 12:45:5[illegible]
Job time : 1[illegible] secs
```

Title: US-09-423-100-2
Perfect score: 470

Database : SwissProt_40:*

SUMMARIES

Result No.	Match Length	ID	Description
1		SOMA_HUMAN	homo sapien
2		SOMA_MACMU	macaca mula
3	217	SOMA_PANTR	P58756 pan troglod
4		SOMA_SAIBB	saimiri bol
		SOMA_CALJA	
		SOMA_PANTR	pan troglod
		SOMA_HUMAN	homo sapien
		SOMA_MACMU	
		PLL_HUMAN	homo sapien
		SOMA_MESAU	mesocricet
		SOMA_BALA	balaenopter
		SOMA_MOUSE	mus muscul
		SOMA_HORSE	equus cab
		SOMA_RABIT	oryctolagus
		SOMA_RAT	rattus
		SOMA_DAUSE	
		SOMA_NYC	nycticebus
		SOMA_LOXAF	loxodonta a
		SOMA_CANFA	canis famil
		SOMA_FELCA	felis silve
		SOMA_PIG	sus scrofa
		SOMA_MUSVI	mustela vis
		SOMA_CAMDR	
		SOMA_VULVU	vulpes
		SOMA_MINK	
		SOMA_TRIVU	trichosurus
		SOMA_BOVIN	bos taurus
		SOMA_CERE	cervus elap
	217	SOMA_SHEEP	ovis aries
		SOMA_BUBBU	bubalus bub
		SOMA_MELCA	
		SOMA_DAMDA	dama dama
		SOMA_SICA	

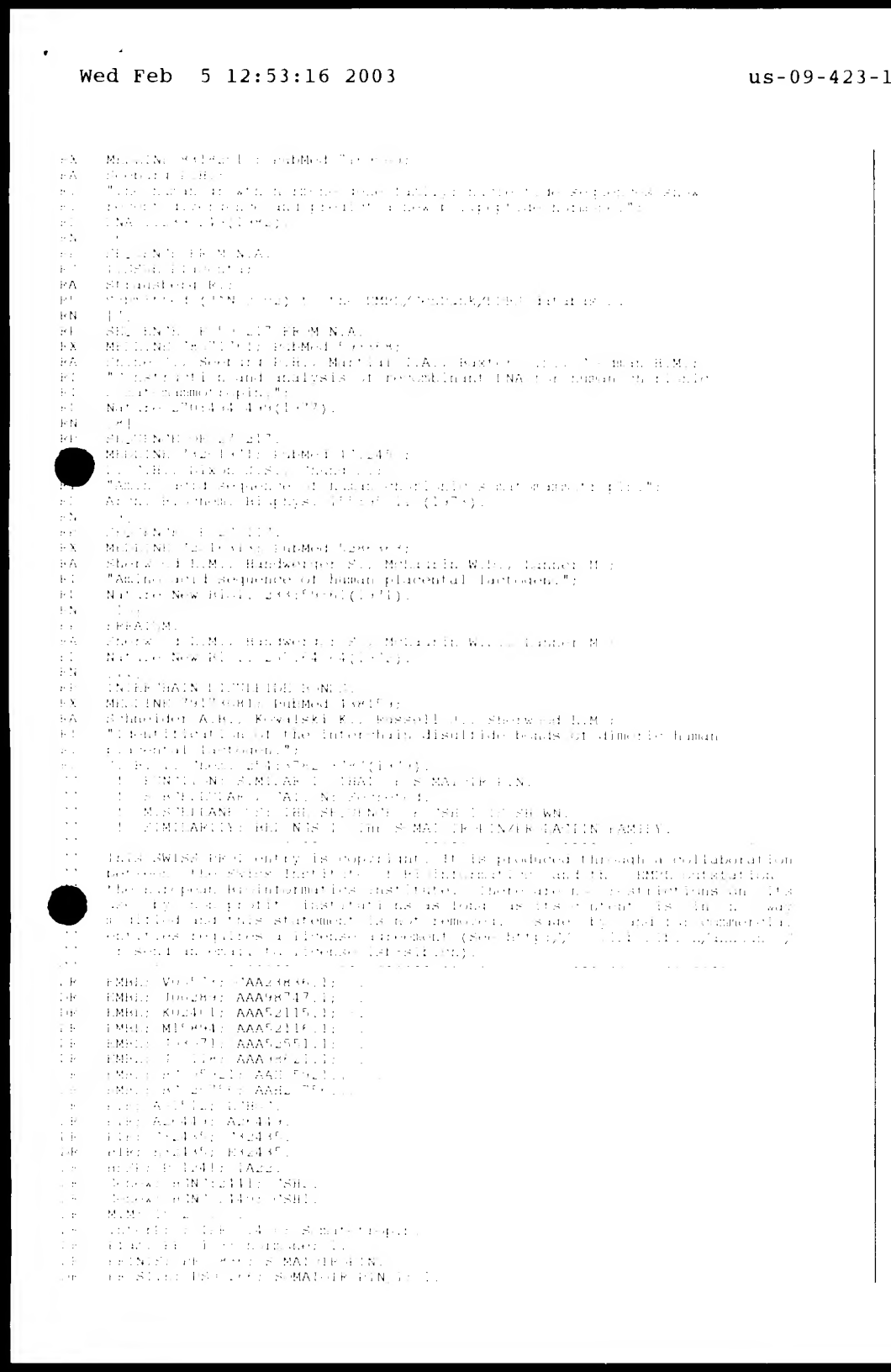
Result No.	ID
40	SOMA_XENLA
42	SOMA_RHOMA
44	SOMA_XENLA

ALIGNMENTS

RESULT 1
SOMA_HUMAN
ID SOMA_HUMAN
DE Somatotropin precursor (Growth hormone) (GH) (GH-N)
DE hormone)
GN GH1
OS Homo sapiens (Human).
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE
RA Roskam W., Rougeon F.;
RT "Molecular cloning and nucleotide sequence of the human growth
RL Nucleic Acids Res.
RN [2]
RP SEQUENCE FROM N.A.
RA Martial J.A., Hallewell R.A., Baxter J.D., Goodman H.M.;
RT "Human growth hormone: complementary DNA cloning and expression
RT bacteria.";
RL Science
RN [3]
RP SEQUENCE FROM N.A.
RT "Human growth hormone DNA sequence and mRNA structure
RT alternative splicing.";
RL Nucleic Acids Res.
RN [4]
RP SEQUENCE FROM N.A.
RL DNA 1:239-249(1982).
RN [5]
RP SEQUENCE FROM N.A.
RN [6]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary;
RT "A novel gene expressed in human pituitary.";
RL Submitted

```
RN   [7]
RT   SEQUENCE FROM N.A. (ISOFORM 4).
RC   TISSUE=Fetal brain;
RA   MEDLINE=94134761; PubMed=7507961;
RA   Barr V.A., Bell G.I., Nelson J.S., Bonifacino J.S.;
RA   "Characterization of the human adaptor protein-1 sigma subunit
RA   (clathrin assembly protein sigma-2) isoforms: genomic structure and
RA   expression.";
RA   J. Biol. Chem. 269:1863-1869(1994).
RN   [8]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96012361; PubMed=8521361;
RA   Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA   Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA   Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA   Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,
RA   Murakami K., Yasuda T., Iwayanagi T., Wada M., Takahashi Y.,
RA   Nakagawa H., Ikuta K., Nakajima K., Yamashita H., Ishikawa H.,
RA   Suto K., Hirakawa M., Hiraoka S., Saito R., Shibata M., Tanaka N.,
RA   Hatsuse H., Abe T., Kikkawa E., Minami-Ishii N., Ohara O.,
RA   Nomura N., Isogai T., Sugano S., Nakai K., Sakamoto S., Yoshinaga K.,
RA   Ishizaki J., Yoshida K., Kikuchi Y., Isogai Z., Yamamoto K.,
RA   Ishibashi K., Toyoda A., Ishii K., Adachi C., Kato M., Saito K.,
RA   Yamaguchi I., Kobayashi M., Miyajima N., Okumura K., Nomaguchi H.,
RA   Isaji S., Takahashi H., Yamada T., Kawakami Y., Tanaka J.,
RA   Totsuka A., Wakisaka A., Sugimoto M., Morishita R., Kim S.,
RA   Iwamatsu A., Kondo H., Yoda Y., Takahashi M., Matsuda H.,
RA   Kanehisa M., Kikuchi S., Saito T., Muramatsu M.;
RT   "Sequence analysis of 40,150 human cDNA clones (KIAA0001-KIAA0500)
RT   obtained from clones with relatively short 3' untranslated regions.";
RL   DNA Res. 2:167-174(1995).
RN   [9]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97289502; PubMed=9184204;
RA   Li C.H., Dixon J.S., Liu W.K.;
RT   "Human pituitary growth hormone, XX. Its primary structure and
RT   evolution.";
RL   Acta Biochim. Biophys. Sin. 29:159-165(1997).
RN   [10]
RP   REVISION.
RX   MEDLINE=96296454; PubMed=8764854;
RA   Powell J.A., Fuqua J.S., Penny J.E.;
RT   "Sequence comparison of human pituitary growth hormone, human
RT   chorionic somatomammotropin, and ovine pituitary growth and
RT   lactogenic hormones.";
RL   Int. J. Pept. Protein Res. 47:281-287(1996).
RN   [11]
RP   SEQUENCE OF 27-217 AND REVISIONS.
RX   MEDLINE=72115035; PubMed=5111622;
RA   Li C.H., Dixon J.S.;
RT   "Human pituitary growth hormone. XX. The primary structure of the
RT   hormone.";
RL   Arch. Biochem. Biophys. 151:334-349(1972).
RN   [12]
RP   REVISION.
RX   MEDLINE=96296454; PubMed=8764854;
RA   Niall H.D.;
RT   "Revised primary structure for human growth hormone.";
RL   Nature New Biol. 230:90-91(1971).
RN   [13]
RP   REVISIONS OF 1-26 AND 167-191.
RX   MEDLINE=71179099; PubMed=5279528;
RA   Niall H.D., Keutmann H.T., Sauer R., Rosenblatt M.,
RA   Deftos L., Potts J.T. Jr.;
RT   "Sequences of pituitary and placental human growth hormones:
RT   evolution of a polypeptide hormone by successive duplications.";
RL   Proc. Natl. Acad. Sci. U.S.A. 68:866-869(1971).
RN   [14]
RP   REVISION.
RA   Niall H.D.;
RT   "The versatility of the human growth hormones.";
RL   (In) Salhanick H. (eds.);
RL   Metabolic and Clinical Aspects of Human Growth Hormone.
RL   pp. 1-9, Alpha Omega Alpha Press, Cardiff (1972).
RN   [15]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RX   MEDLINE=91135161; PubMed=2001224;
RA   Lupski J.R., Montes de Oca-Luna R., Slaugenhaupt S., Pentao L.,
RA   Guzetta V., Wise C.A., Carey K.D., Patel P.I., Quan F., Matsunami N.,
RA   Weleber R.G., Stringham H.M., Tsui L.-C., Broman K., Haines J.L.,
RA   Torbergsen T., Berg K., Gilliam T.C., Wadelius C., Kidd K.K.,
RA   Roderick T.H., Seldin M.F., Morton N.E., Motulsky A.G., Patil N.,
RA   Klein D., O'Brien S.J., Magenis R.E., Zackai E.H., Emanuel B.,
RA   Budarf M., Driscoll D.A., Wang E., Wang P., Wang Z., Bray-Ward P.,
RA   Ward D.C., Weissenbach J., Fox M., Devoto M., Obata F., Kere J.,
RA   Nakamura Y., Hsieh C.-L., Olsen A.S., Buetow K., Fan Y.-S.,
RA   Taillon-Miller P., Brzustowicz L.M., Gilliam T.C., Kidd J.R.,
RA   Kidd K.K., Daiger S.P., Pericak-Vance M.A.;
RT   "Genetic mapping of the Charcot-Marie-Tooth disease variant in
RT   chromosome region 17p11.2.";
RL   Cell 66:219-232(1991).
RN   [16]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RX   MEDLINE=96012361; PubMed=8521361;
RA   Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA   Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA   Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA   Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,
RA   Murakami K., Yasuda T., Iwayanagi T., Wada M., Takahashi Y.,
RA   Nakagawa H., Ikuta K., Nakajima K., Yamashita H., Ishikawa H.,
RA   Suto K., Hirakawa M., Hiraoka S., Saito R., Shibata M., Tanaka N.,
RA   Hatsuse H., Abe T., Kikkawa E., Minami-Ishii N., Ohara O.,
RA   Nomura N., Isogai T., Sugano S., Nakai K., Sakamoto S., Yoshinaga K.,
RA   Ishizaki J., Yoshida K., Kikuchi Y., Isogai Z., Yamamoto K.,
RA   Ishibashi K., Toyoda A., Ishii K., Adachi C., Kato M., Saito K.,
RA   Yamaguchi I., Kobayashi M., Miyajima N., Okumura K., Nomaguchi H.,
RA   Isaji S., Takahashi H., Yamada T., Kawakami Y., Tanaka J.,
RA   Totsuka A., Wakisaka A., Sugimoto M., Morishita R., Kim S.,
RA   Iwamatsu A., Kondo H., Yoda Y., Takahashi M., Matsuda H.,
RA   Kanehisa M., Kikuchi S., Saito T., Muramatsu M.;
RT   "Sequence analysis of 40,150 human cDNA clones (KIAA0001-KIAA0500)
RT   obtained from clones with relatively short 3' untranslated regions.";
RL   DNA Res. 2:167-174(1995).
RN   [17]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96296454; PubMed=8764854;
RA   Powell J.A., Fuqua J.S., Penny J.E.;
RT   "Sequence comparison of human pituitary growth hormone, human
RT   chorionic somatomammotropin, and ovine pituitary growth and
RT   lactogenic hormones.";
RL   Int. J. Pept. Protein Res. 47:281-287(1996).
RN   [18]
RP   REVIEW.
RX   MEDLINE=96317812; PubMed=8782801;
RA   Baumann G.;
RT   "Growth hormone heterogeneity: genes, isohormones, variants, and
RT   binding proteins.";
RL   Endocr. Rev. 17:237-269(1996).
RN   [19]
RP   SPECIFIC PEPTIDE MODELING.
RX   MEDLINE=88166744; PubMed=3127113;
RA   Cohen F.E., Kuntz I.D.;
RT   "Prediction of the three-dimensional structure of human growth
RT   hormone.";
RL   Proteins 2:162-166(1987).
RN   [20]
RP   X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
RX   MEDLINE=96196466; PubMed=8601825;
RA   de Vos A.M., Ultsch M., Kossiakoff A.A.;
RT   "Human growth hormone and extracellular domain of its receptor:
RT   crystal structure of the complex.";
RL   Science 255:306-312(1992).
RN   [21]
RP   X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RX   MEDLINE=95079462; PubMed=7984111;
RA   Somers W., Ultsch M., de Vos A.M., Kossiakoff A.A.;
RT   "The X-ray structure of a growth hormone-prolactin receptor complex.";
RL   Nature 372:478-481(1994).
RN   [22]
RP   X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RA   Chantalat L., Chirgadze N.Y., Jones N., Korber F., Nayeem A.,
RA   Pavlovsk A.G., Wlodawer A.;
RT   "The crystal structure of wild-type growth hormone at 2.5 A
RT   resolution.";
RL   Protein Sequ. Data Anal. 5:43-51(1992).
RN   [23]
RP   X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
RX   MEDLINE=97116025; PubMed=8918476;
RA   Sundstrom M., Lundqvist T., Rodin A., Lundqvist H., Mattsson O.,
RA   Norstedt G.;
RT   "Crystal structure of an oncogenic mutant of human growth hormone
RT   G120R, incomplex with its receptor at 2.6 A resolution.";
RL   J. Biol. Chem. 271:32967-32970(1996).
CC   -!- FUNCTION: Stimulates growth and cell reproduction. It is
CC       noted it stimulates bone growth as well as protein synthesis.
CC       Other tissues to which it stimulates both the
CC       differentiation and proliferation of myoblasts, stimulates
CC       amino acid uptake and protein synthesis in muscle and other
CC       tissues.
CC   -!- SUBUNIT: Monomer. Also exists in a dimeric and tetrameric
CC       forms in the pituitary, as well as in soluble complexes with
CC       hGHBP or with the alpha2-macroglobulin complex.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- ALTERNATIVE PRODUCTS: At least 4 isoforms, 1/2a, 1/2b (22 kDa),
CC       2/3a, 2/3b (20 kDa) variants. 1 and 2 are produced by alternative
CC       splicing.
CC   -!- DISEASE: DEFECTS IN GH1 ARE A CAUSE OF PITUITARY DWARFISM IA AND
CC       IV.
CC   -!- PHARMACEUTICAL: Available under the names Nutropin or Protropin
CC       (Genentech), Norditropin (Novo Nordisk), Saizen (Serono),
CC       Genotropin (Pharmacia & Upjohn), Humatrope (Lilly), Omnitrope
CC       (Sandoz) or Valtropin (Biopartners). Used for the treatment of growth disorders and for
CC       human growth hormone deficiency and for
CC       Turner's syndrome.
```

```
FT   CHAIN        27    217       GROWTH HORMONE VARIANT.
FT   CONFLICT     ?     ?         BY SIMILARITY.
FT   CONFLICT     ?     ?         BY SIMILARITY.
FT   VARIANT      169   169       N->K,NKEE (5,'NA',3,3) (1 GENTIAL).
FT   VARSPLIC     154   217       HEEGSERELQLLNGVSKEPHKSRNEGALFLKNYIYGH
FT                                KMERVETFLEVVGSVVGKYSF    VGVAGSLNSVIAP
FT                                IAGNGAREKHFILDSDVALRYRSTPSLVNLHMGGL
FT                                PGHEDKKMNGAGGQGQLQKAVLQAALDLEAFHHHGT
FT                                VGLSVLGLV (IN REF. 2).
FT   CONFLICT     ?     ?         ? - ? (IN REF. 2).
SQ   SEQUENCE   217 AA;  24889 MW;  7B9247D6B3CEF0C0 CRC64;

  Query Match             89.8%;  Score 422;  DB 1;  Length 217;
  Best Local Similarity   92.6%;  Pred. No. e-56;
  Matches  201;  Conservative    5;  Mismatches    8;  Indels    3;  Gaps    3;

Qy       2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPT 61
           |||||||||||||||||| |||||| |||||||||| ||||||||||||||| |||||||
Db      27 FPTIPLSRLFDNAMLRAHRLYQLAYDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPT 86

Qy      62 FSNREETQQKSNLELLRISLLLIQSWLEPVQ 92
           |||||||||||||||||||||||||||||||
Db      87 FSNREETQQKSNLELLRISLLLIQSWLEPVQ 117


RESULT 2
S_M_MAMM
ID   S_M_MAMM       STANDARD;      PRT:   217 AA.
AC   ;
DT   01-NOV-1994 (Rel. 31, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Growth hormone variant precursor (GHV) (Placental prolactin-like
DE   hormone) (Growth hormone 2).
GN   GHL.
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Soboleva I.V.;
RT   "Submitted (JAN-1994) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=95008424; PubMed=7961747;
RA   Miller W.L., Eberhardt N.L., Rotwein P., Bell G.I., Boime I.;
RT   "Structure of four growth hormone/chorionic somatomammotropin genes.";
RL   DNA Cell Biol. 13:1-15(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=95008424; PubMed=7961747;
RA   Miller W.L., Eberhardt N.L., Rotwein P., Bell G.I., Boime I.;
RT   "Structure of four growth hormone/chorionic somatomammotropin genes.";
RL   DNA Cell Biol. 13:1-15(1994).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=95008424; PubMed=7961747;
RA   Miller W.L., Eberhardt N.L., Rotwein P., Bell G.I., Boime I.;
RT   "Structure of four growth hormone/chorionic somatomammotropin genes.";
RL   DNA Cell Biol. 13:1-15(1994).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=95008424; PubMed=7961747;
RA   Miller W.L., Eberhardt N.L., Rotwein P., Bell G.I., Boime I.;
RT   "Structure of four growth hormone/chorionic somatomammotropin genes.";
RL   DNA Cell Biol. 13:1-15(1994).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=95008424; PubMed=7961747;
RA   Miller W.L., Eberhardt N.L., Rotwein P., Bell G.I., Boime I.;
RT   "Structure of four growth hormone/chorionic somatomammotropin genes.";
RL   DNA Cell Biol. 13:1-15(1994).
CC   -!- FUNCTION: Plays an important role in growth control. Its major
CC       role in stimulating body growth is to stimulate the liver and
CC       other tissues to secrete IGF-I. It stimulates both the
CC       differentiation and proliferation of myoblasts. It also stimulates
CC       amino acid uptake and protein synthesis in muscle and other
CC       tissues.
CC   -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC   -!- TISSUE SPECIFICITY: Expressed in the placenta.
CC   -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   PIR; S22255; AAA49181; -.
DR   EMBL; L16555; AAA20903; -.

RESULT 3
PEL_HUMAN
ID   PEL_HUMAN      STANDARD;      PRT:   217 AA.
AC   P01241;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-APR-1990 (Rel. 08, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Lactogen precursor (Chorionammotropin) (Chorionic somatomammotropin).
GN   CSH1 AND CSH2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (CLONE CSH1).
RX   MEDLINE=85099129; PubMed=6299302;
RA   Seeburg P.H., Barta A., Baxter J.D., Shine J., Sherwood L.M.;
RT   "Adjacent human growth hormone and chorionic somatomammotropin genes:
RT   both genes are preceded by Alu family sequences.";
RL   DNA 3:37-44(1984).
RN   [2]
RP   SEQUENCE FROM N.A. (CLONE CSH1).
RX   MEDLINE=87172475; PubMed=3031031;
RA   Barta A., Richards R.I., Baxter J.D., Shine J., Seeburg P.H.;
RT   "The human growth hormone gene contains structures, evidence of
RT   allelic variations.";
RL   DNA 6:515-527(1987).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87172475; PubMed=3031031;
RA   Barta A., Richards R.I., Baxter J.D., Shine J., Seeburg P.H.;
RT   "The human growth hormone gene contains structures, evidence of
RT   allelic variations.";
RL   DNA 6:515-527(1987).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87172475; PubMed=3031031;
RA   Barta A., Richards R.I., Baxter J.D., Shine J., Seeburg P.H.;
RT   "The human growth hormone gene contains structures, evidence of
RT   allelic variations.";
RL   DNA 6:515-527(1987).
RN   [5]
RP   SEQUENCE FROM N.A. (GENES CSH1 AND CSH2).
RX   MEDLINE=87277; PubMed=2431026;
RA   Chen E.Y., Liao Y.C., Smith D.H., Barrera-Saldana H.A., Gelinas R.E.,
RA   Seeburg P.H.;
RT   "The human growth hormone gene family: nonidentical genes expressed in the
RT   pituitary and placenta.";
RL   DNA 8:239-249(1989).
RN   [6]
RP   SEQUENCE.
```



```
DE   Pfam; PF00103; hormone; 1.
DR   PRINTS; PR00836; SOMATOTROPIN.
DR   ProSite; PS00266; SOMATOTROPIN_1; 1.
DR   ProSite; PS00267; SOMATOTROPIN_2; 1.
KW   Hormone; Pituitary; Signal.
FT   SIGNAL       1     26       BY SIMILARITY.
FT   CHAIN       27    216       SOMATOTROPIN.
FT   DISULFID    76    204       BY SIMILARITY.
FT   DISULFID   206    214       BY SIMILARITY.
SQ   SEQUENCE   216 AA;  24696 MW;  3EC9CD12AD6F1196 CRC64;

  Query Match             66.1%;  Score 410.5;  DB 1;  Length 216;
  Best Local Similarity   67.0%;  Pred. No. 6e-26;
  Matches   61;  Conservative   14;  Mismatches   15;  Indels    1;  Gaps    1;

Qy       2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPTPS 61
           ||  |||| || ||||  |||||| ||||||| |||| |||| ||| ||  |||||||  |||
Db      27 FPAMSVSSLFANAVLRAQHLAQLAADTFKEFERTYIPEGQRYSIQNAQAAFCFSETIPAPT 86

Qy      62 PSNREETQQKSNLELLRISLLLIQSWLEPVQ 92
           |  |||| |||| |||||||||| |||
Db      87 PTGKDEAQQKSDMELLRFSLLLIQSWLGPVQ 117

RESULT 11
SOMA_BALAC
ID   SOMA_BALAC     STANDARD;      PRT;   216 AA.
AC   P03892;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Somatotropin (Growth hormone).
GN   GH1.
OS   Balaenoptera borealis (Sei whale).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Cetacea; Mysticeti;
OC   Balaenopteridae; Balaenoptera.
OX   NCBI_TaxID=9774;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=83028669; PubMed=7115913;
RA   Tudawe N.A., Fankboner P.A., Bolaffi A.A., Wallpora I.A.;
RT   "Amino acid sequence of bridled somatotropin.";
RL   Biokhimiia 47:1059-1069(1982).
RN   [2]
RP   PRELIMINARY PARTIAL SEQUENCE.
RA   Wallpora I.A., Bolaffi A.A., Tudawe N.A.;
RT   "Structural studies of tryptic peptides from large cyanogen bromide
RT   fragments of sei whale (Balaenoptera borealis) somatotropin.";
RL   Bioorg. Khim. 4:1589-1599(1978).
CC   -!- FUNCTION: Plays an important role in growth control. Its major
CC       role in stimulating body growth is to stimulate the liver and
CC       other tissues to secrete IGF-1. It stimulates both the
CC       differentiation and proliferation of myoblasts. It also stimulates
CC       amino acid uptake and protein synthesis in muscle and other
CC       tissues.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN FAMILY.
DR   PIR; PN0140; PN0140.
DR   PIR; JN0387; JN0387.
DR   HSSP; P01241; 1AXI.
DR   InterPro; IPR001400; Somatotropin.
DR   Pfam; PF00103; hormone; 1.
DR   PRINTS; PR00836; SOMATOTROPIN.
DR   ProSite; PS00266; SOMATOTROPIN_1; 1.
DR   ProSite; PS00267; SOMATOTROPIN_2; 1.
KW   Hormone; Pituitary.
FT   DISULFID    52    164       BY SIMILARITY.
FT   DISULFID   190    198       BY SIMILARITY.
SQ   SEQUENCE   216 AA;  24665 MW;  69FB8F63B19A7516 CRC64;

  Query Match             65.4%;  Score 407.5;  DB 1;  Length 216;
  Best Local Similarity   67.0%;  Pred. No. 1.1e-25;

  Matches   61;  Conservative   14;  Mismatches   15;  Indels    1;  Gaps    1;

Qy       2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPTPS 61
           ||  |||| || ||||  |||||| ||||||| |||| |||| ||| ||  |||||||  |||
Db      27 FPAMSVSSLFANAVLRAQHLAQLAADTFKEFERTYIPEGQRYSIQNAQAAFCFSETIPAPT 86

Qy      62 PSNREETQQKSNLELLRISLLLIQSWLEPVQ 92
           |  |||| |||| |||||||||| |||
Db      87 PTGKDEAQQKSDMELLRFSLLLIQSWLGPVQ 117

RESULT 12
SOMA_MOUSE
ID   SOMA_MOUSE     STANDARD;      PRT;   216 AA.
AC   P06880;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-JAN-1988 (Rel. 06, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Somatotropin precursor (Growth hormone).
GN   GH1 OR GH.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Gaudray G., Devaux F., Tassin J., Kerbiriou D., Moreau-Gachelin F.;
RA   Linzer D.I.H., Talamantes F.;
RT   "Nucleotide sequence of mouse prolactin and growth hormone mRNAs and
RT   expression of these mRNAs during pregnancy.";
RL   J. Biol. Chem. 260:9574-9579(1985).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN PZIDU; TISSUE=Liver;
RX   MEDLINE=96193877; PubMed=8599116;
RA   Bar P., Meyer S., Gyapay G., Weissenbach J., Petzsch M.;
RT   "Structure of the growth hormone encoding gene and its promoter in
RT   mice.";
RL   Gene 169:209-213(1996).
CC   -!- FUNCTION: Plays an important role in growth control. Its major
CC       role in stimulating body growth is to stimulate the liver and
CC       other tissues to secrete IGF-1. It stimulates both the
CC       differentiation and proliferation of myoblasts. It also stimulates
CC       amino acid uptake and protein synthesis in muscle and other
CC       tissues.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is not
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M22541; CAA26650.1;
DR   EMBL; Z22555; CAA80658.1;
DR   PIR; B23911; SIMS.
DR   HSSP; P01246; 1BST.
DR   MGD; MGI:95787; Gh.
DR   InterPro; IPR001400; Somatotropin.
DR   Pfam; PF00103; hormone; 1.
DR   PRINTS; PR00836; SOMATOTROPIN.
DR   ProSite; PS00266; SOMATOTROPIN_1; 1.
DR   ProSite; PS00267; SOMATOTROPIN_2; 1.
KW   Hormone; Pituitary; Signal.
FT   SIGNAL       1     26       BY SIMILARITY.
FT   CHAIN       27    216       SOMATOTROPIN.
FT   DISULFID    76    204       BY SIMILARITY.
FT   DISULFID   206    214       BY SIMILARITY.
SQ   SEQUENCE   216 AA;  24716 MW;  0A66AA4A82D6F5F7 CRC64;
```

```
  Query Match             74.8%;  Score 554.5;  DB 1;  Length 216;
  Best Local Similarity   64.9%;  Pred. No. 2.6e-56;
  Matches  99;  Conservative   14;  Mismatches   17;  Indels    1;  Gaps    1;

Qy       2 PPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPT 61
           | ||||  ||  |||||  |||| |||||| |||| |||  ||  | |  ||||||
Db      27 FPAMSLSSLFANAVLRAQHLHQLAADTFKEFERAYIPEGQRYSIQNAQAAFCFSETIPA 86

Qy      62 PSNKEETQQKSNLELLRISLLLIQSWLEPV 92
           || ||| |||||||| |||||||| |||
Db      87 PTGKDEAQQRSDMELLRFSLLLIQSWLGPV 117


RESULT 14
SOMA_RABIT
ID   SOMA_RABIT     STANDARD;      PRT:   216 AA.
AC   P46407;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Somatotropin precursor (Growth hormone).
GN   GH1.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
X    M76115; Ia410.1996;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=New Zealand white;
RX   MEDLINE=96111664; PubMed=7501127;
RA   Wallis M.; Wallis M.;
RT   "Cloning and characterisation of the rabbit growth hormone
RT   gene.";
RL   Gene 166:253-256(1995).
CC   -!- FUNCTION: Plays an important role in growth. Controls
CC       role in stimulating body growth is to stimulate the liver and
CC       other tissues to secrete IGF-1. It stimulates both the
CC       differentiation and proliferation of myoblasts. It also stimulates
CC       amino acid uptake and protein synthesis in muscle and other
CC       tissues.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN FAMILY.
CC   ----------  ----------  ----------  ----------  ----------  ----------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------  ----------  ----------  ----------  ----------  ----------
DR   EMBL; D01229; AAA31027.1; -.
DR   PIR; A01514; STHR.
DR   HSSP; P01246; 1HST.
DR   InterPro; IPR001400; Somatotropin.
DR   Pfam; PF00103; hormone; 1.
DR   PRINTS; PR00836; SOMATOTROPIN.
DR   PROSITE; PS00264; SOMATOTROPIN_1; 1.
DR   PROSITE; PS00338; SOMATOTROPIN_2; 1.
KW   Hormone; Pituitary; Signal.
FT   SIGNAL        1     27
FT   CHAIN        28    216       SOMATOTROPIN.
FT   DISULFID     53    165
FT   DISULFID     82    189
FT   DISULFID    182    207
SQ   SEQUENCE   216 AA;  24116 MW;  F7B8C1C8C1E1A7A0 CRC32;

  Query Match             74.8%;  Score 554.5;  DB 1;  Length 216;
  Best Local Similarity   64.9%;  Pred. No. 2.6e-56;
  Matches  99;  Conservative   14;  Mismatches   17;  Indels    1;  Gaps    1;

Qy       2 PPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPT 61
           | ||||  ||  |||||  |||| |||||| |||| |||  ||  | |  ||||||
Db      27 FPAMSLSSLFANAVLRAQHLHQLAADTFKEFERAYIPEGQRYSIQNAQAAFCFSETIPA 86

Qy      62 PSNKEETQQKSNLELLRISLLLIQSWLEPV 92
           || ||| |||||||| |||||||| |||
Db      87 PTGKDEAQQRSDMELLRFSLLLIQSWLGPV 117


RESULT 14
SOMA_RABIT
ID   SOMA_RABIT     STANDARD;      PRT:   216 AA.
AC   P46407;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Somatotropin precursor (Growth hormone).
GN   GH1.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
X    M76115; Ia410.1996;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=New Zealand white;
RX   MEDLINE=96111664; PubMed=7501127;
RA   Wallis M.; Wallis M.;
RT   "Cloning and characterisation of the rabbit growth hormone
RT   gene.";
RL   Gene 166:253-256(1995).
CC   -!- FUNCTION: Plays an important role in growth. Controls
CC       role in stimulating body growth is to stimulate the liver and
CC       other tissues to secrete IGF-1. It stimulates both the
CC       differentiation and proliferation of myoblasts. It also stimulates
CC       amino acid uptake and protein synthesis in muscle and other
CC       tissues.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN FAMILY.
CC   ----------  ----------  ----------  ----------  ----------  ----------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------  ----------  ----------  ----------  ----------  ----------
DR   EMBL; X58124; CAA41408.1; -.
DR   HSSP; P01241; 1HST.
DR   InterPro; IPR001400; Somatotropin.
DR   Pfam; PF00103; hormone; 1.
```

```
DR   PRINTS; PR00836; SOMATOTROPIN.
DR   PROSITE; PS00266; SOMATOTROPIN_1; 1.
DR   PROSITE; PS00267, SOMATOTROPIN_2; 1.
KW   Hormone; Pituitary; Signal.
FT   SIGNAL          1     26       POTENTIAL.
FT   CHAIN          27    216       SOMATOTROPIN.
FT   DISULFID       78    189       BY SIMILARITY.
FT   DISULFID      206    214       BY SIMILARITY.
SQ   SEQUENCE   216 AA;  24158 MW;  4B2F1F7181994B25 CRC64;

  Query Match             64.4%;  Score 382.5;  DB 1;  Length 216;
  Best Local Similarity   64.8%;  Prod. No. 4.3e-25;
  Matches   59;  Conservative   34;  Mismatches   17;  Indels    1;  Gaps    1;

Qy       2 PFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPT 61
           ||  ||| || |:|||  |||:| |||:||| |||| |:|| :|: | :  |||:||
Db      27 PPAMDLKFLQDNAQLAAIHILHSLKNYIEFLNVLTGKWEDSNTMMESLVMIGEQLDKTIA 86

Qy      62 PSNREETQQKSNLELLRISLLLIQSWLEPVQ 92
           |: ::| |||:|||| |||||||| |||
Db      86 PIGKEDNDSPETMEFQFQIAEQIRGQLERVQ 116


RESULT 15
SOMA_BAT
ID   SOMA_BAT        STANDARD;       PRT;   216 AA.
AC   P01244;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Somatotropin precursor (Growth hormone).
GN   GH1 OR GH.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=82059626; PubMed=6272224;
RA   Page R.B., Smith S., Goodman H.M.;
RT   "DNA sequence of the rat growth hormone gene: location of the 5'
RT   terminus of the growth hormone mRNA and identification of an internal
RT   transposon like element.";
RL   Nucleic Acids Res. 9:2882-2901(1981).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=78071760; PubMed=149105;
RA   Seeburg P.H., Shine J., Martial J.A., Baxter J.D., Goodman H.M.;
RT   "Nucleotide sequence and amplification in bacteria of structural gene
RT   for rat growth hormone.";
RL   Nature 270:486-494(1977).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=82066155; PubMed=6946433;
RA   Barta A., Richards R.I., Baxter J.D., Shine J.;
RT   "Primary structure and evolution of rat growth hormone gene.";
RL   Proc. Natl. Acad. Sci. U.S.A. 78:4867-4871(1981).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley;
RX   MEDLINE=96056684; PubMed=8521139;
RA   Eskay M.M., Weigent D.A.;
RT   "Cloning and nucleotide sequencing of rat lymphocyte growth hormone
RT   cDNA.";
RL   Neuroimmunomodulation 2:108-114(1995).
CC   -!- FUNCTION: Plays an important role in growth control. Its major
CC       role in stimulating body growth is to stimulate the liver and
CC       other tissues to secrete IGF-1. It stimulates both the
CC       differentiation and proliferation of myoblasts. It also stimulates
CC       amino acid uptake and protein synthesis in muscle and other
CC       tissues.
CC   -!- SUBCELLULAR LOCATION: Secreted.

1  pfptiplsrlfdnamlrahrlhqlafdtyqefeeayipkeqkysflqnpqtslcfsesiptp 60
  ||  ||| || |:|||  |||:| |||:||| |||| |:|| :|: | :  |||:||  |
2  ppamdlkflqdnaqlaaihilhslknyiefleavlpeekkysflqnpqtslcfsesiptp 61

61  psnreetqqksnlellrislllqswlepvq 92
  |: ::| |||:|||| |||||||| |||
62  pigkednespetmefqfqiaeqirgqlervq 92

Search completed: February  5, 2003, 12:14:57
Job time : 2.24471 secs
```



```
DR   EMBL; V01267; CAA24547.1; -.
DR   EMBL; V01268; CAA24548.1; -.
DR   EMBL; V01270; CAA24549.1; -.
DR   EMBL; U67277; AAB04025.1; -.
DR   PIR; A01511; GTRT.
DR   HSSP; P01246; 1BST.
DR   InterPro; IPR006612; Somato_tropin.
DR   Pfam; PF00103; hormone; 1.
DR   PRINTS; PR00836; SOMATOTROPIN.
DR   PROSITE; PS00266; SOMATOTROPIN_1; 1.
DR   PROSITE; PS00267; SOMATOTROPIN_2; 1.
KW   Hormone; Pituitary; Signal.
FT   SIGNAL          1     26
FT   CHAIN          27    216       SOMATOTROPIN.
FT   DISULFID       78    189       BY SIMILARITY.
FT   DISULFID      206    214       BY SIMILARITY.
FT   CONFLICT       27     27       F    L (IN REF. 2 AND 4).
SQ   SEQUENCE   216 AA;  24658 MW;  [illegible] CRC64;

  Query Match             64.4%;  Score 382.5;  DB 1;  Length 216;
  Best Local Similarity   64.8%;  Prod. No. 4.3e-25;
  Matches   59;  Conservative   34;  Mismatches   17;  Indels    1;  Gaps    1;

Qy       2 PFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPT 61
           ||  ||| || |:|||  |||:| |||:||| |||| |:|| :|: | :  |||:||
Db      27 PPAMDLKFLQDNAQLAAIHILHSLKNYIEFLNVLTGKWEDSNTMMESLVMIGEQLDKTIA 86

Qy      62 PSNREETQQKSNLELLRISLLLIQSWLEPVQ 92
           |: ::| |||:|||| |||||||| |||
Db      86 PIGKEDNDSPETMEFQFQIAEQIRGQLERVQ 116


Search completed: February  5, 2003, 12:14:57
Job time : 2.24471 secs
```

Run on: February 5, 2003, 12:42:21 ; Search time 0.072 Seconds
(without alignments)
7.127 Million cell updates/sec

Title: US-09-423-100-2
Perfect score: 470
Sequence: 1 MEEIIPLSPEENAMLRARR.....NLEIIPCSEIIIQWEEIVQ 472

Scoring table: BLOSUM62
Gap 10.0 , Gap 0.1

Searching: 1,716,466 sequences, 447,117,338 residues

Total number of hits satisfying chosen parameters: 1716466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_synthetic:*
13: sp_unclassified:*
14: sp_virus:*
15: sp_bacteria_2:*
16: sp_archaea_2:*

Prob. No. is the number of results predicted by chance to score as well or better than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

No.	Score	Match	Length	DB	ID	Description
1	459	97.7	217	4	Q9WNE7	Mature alleles mout
2	454	96.6	248	3	Q14111	Homo sapien
3	453	96.4	217	4	Q9WNE8	Mature alleles mout
4	453	96.4	217	4	Q9WNE9	Mature alleles mout
5	451	95.9	217	4	Q14107	Homo sapien
6	428	91.0	217	4	Q9WNE6	Mature alleles mout
7	425.5	90.5	222	4	Q14112	Homo sapien
8	424.5	90.3	217	4	Q9WNE5	Mature alleles mout
9	424.5	90.3	224	4	Q9WNE4	Mature alleles mout
10	423.5	90.1	219	14	Q9YKM4	Avian sarcoma virus
11	422.5	89.8	142	4	Q9WNE3	Mature alleles mout
12	422.5	89.8	219	4	Q9WNE2	Mature alleles mout
13	422.5	89.8	224	4	Q9WNE1	Mature alleles mout
14	422.5	89.8	224	4	Q9WNE0	Mature alleles mout
15	422.5	89.8	224	4	Q9WNE8	Mature alleles mout
16	422.5	89.8	224	4	Q9WNE9	Mature alleles mout
17	421.5	89.6	224	4	Q9WNE6	Mature alleles mout
18	421.5	89.6	224	4	Q9WNE5	Mature alleles mout
19	420	89.2	224	4	Q9WNE4	Mature alleles mout
20	419.5	89.1	224	4	Q9WNE3	Mature alleles mout
21	419.5	89.1	224	4	Q9WNE2	Mature alleles mout
22	419.5	89.1	224	4	Q9WNE1	Mature alleles mout
23	419.5	89.1	224	4	Q9WNE0	Mature alleles mout
24	419.5	89.1	224	4	Q9WNE8	Mature alleles mout
25	419	89.0	224	4	Q9WNE9	Mature alleles mout
26	418.5	88.9	224	4	Q9WNE6	Mature alleles mout
27	418.5	88.9	224	4	Q9WNE5	Mature alleles mout
28	418.5	88.9	224	4	Q9WNE4	Mature alleles mout
29	418.5	88.9	224	4	Q9WNE3	Mature alleles mout
30	418.5	88.9	224	4	Q9WNE2	Mature alleles mout
31	418.5	88.9	224	4	Q9WNE1	Mature alleles mout
32	418.5	88.9	224	4	Q9WNE0	Mature alleles mout
33	418.5	88.9	224	4	Q9WNE8	Mature alleles mout
34	418.5	88.9	224	4	Q9WNE9	Mature alleles mout
35	418.5	88.9	224	4	Q9WNE6	Mature alleles mout
36	418.5	88.9	224	4	Q9WNE5	Mature alleles mout
37	418.5	88.9	224	4	Q9WNE4	Mature alleles mout
38	418.5	88.9	224	4	Q9WNE3	Mature alleles mout
39	418.5	88.9	224	4	Q9WNE2	Mature alleles mout
40	418.5	88.9	224	4	Q9WNE1	Mature alleles mout
41	418.5	88.9	224	4	Q9WNE0	Mature alleles mout
42	418.5	88.9	224	4	Q9WNE8	Mature alleles mout
43	418.5	88.9	224	4	Q9WNE9	Mature alleles mout
44	418.5	88.9	224	4	Q9WNE6	Mature alleles mout
45	418.5	88.9	224	4	Q9WNE5	Mature alleles mout

ALIGNMENTS

RESULT 1
Q9WNE7
ID Q9WNE7 PRELIMINARY; PRT; 217 AA.
AC Q9WNE7;
DT 01-MAR-2000 (TREMBLrel. 13, Created)
DT 01-MAR-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE Mature alleles mouse
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Wong A.L., Hedrick J.L., Satta Y., Takahata N.;
RT "Independent duplication of the gnathostome interferon gamma
RT multigene family.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF154107; AAD51107.1; -.
DR EMBL; AF154108; AAD51108.1; -.
DR EMBL; AF154109; AAD51109.1; -.
DR EMBL; AF154110; AAD51110.1; -.
DR EMBL; AF154111; AAD51111.1; -.
DR EMBL; AF154112; AAD51112.1; -.
DR EMBL; AF154113; AAD51113.1; -.
DR EMBL; AF154114; AAD51114.1; -.
DR EMBL; AF154115; AAD51115.1; -.
SQ SEQUENCE 217 AA; 24111 MW; 5A1B2C3D4E5F6A7B CRC64;

Query Match 97.7%; Score 459; DB 4; Length 217;
Best Local Similarity 97.7%; Pred. No. 1.1e-133;
Matches 459; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 MEEIIPLSPEENAMLRARRLNLAVAKEDLKKVKDFLKANATDVPNWSVDTFNTSLLDL 60
Db 1 MEEIIPLSPEENAMLRARRLNLAVAKEDLKKVKDFLKANATDVPNWSVDTFNTSLLDL 60

Qy 61 KTQLEAKSFDDFEKQFEEETLKSAKEIPSERITEEGQDKPEIQHFKPLKEEQPEVAQSS 120
Db 61 KTQLEAKSFDDFEKQFEEETLKSAKEIPSERITEEGQDKPEIQHFKPLKEEQPEVAQSS 120

Qy 121 VESHSYFTITGLTEDTTQFEAVMKVYNNDKQRDESLILDIIRYSGCSGSATLTGWSVLSG 180
Db 121 VESHSYFTITGLTEDTTQFEAVMKVYNNDKQRDESLILDIIRYSGCSGSATLTGWSVLSG 180

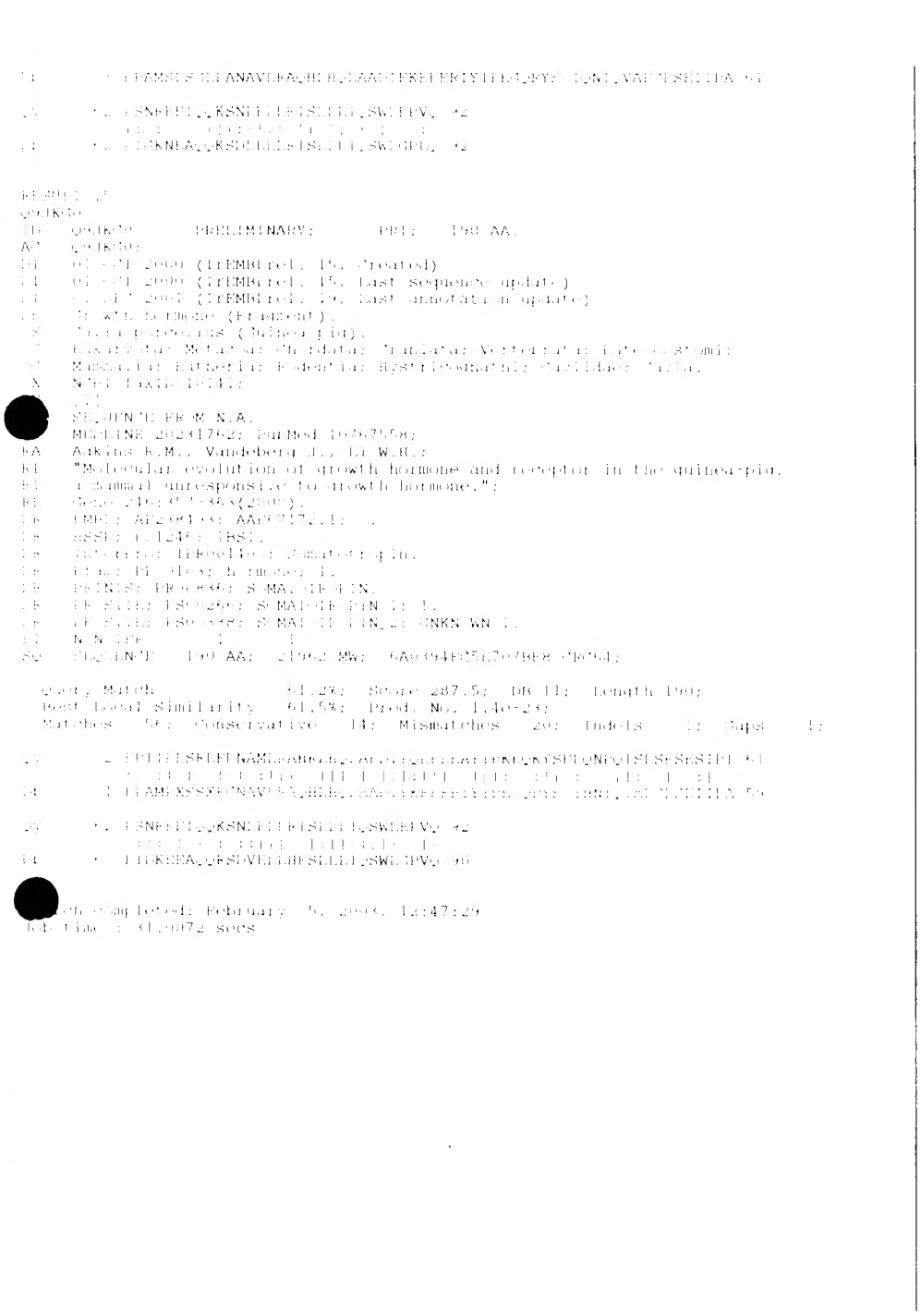
Qy 181 LRSLKLVDVHRHVWLAHRLKEIQDEQNVEHSMSFVKSHLRSEETSLEVSSQVSQFYHKLN 240
Db 181 LRSLKLVDVHRHVWLAHRLKEIQDEQNVEHSMSFVKSHLRSEETSLEVSSQVSQFYHKLN 240

Qy 241 LLAGDKDSSSGAVSAIVQNSQLEQLIGQWELIVQ 274
Db 241 LLAGDKDSSSGAVSAIVQNSQLEQLIGQWELIVQ 274

```
DT   [illegible]-NOV-1999 (TrEMBLrel. [illegible], Created)
DT   [illegible]-NOV-1999 (TrEMBLrel. [illegible], Last sequence update)
DT   [illegible] (TrEMBLrel. [illegible], Last annotation update)
DE   Somatotropin 1 precursor (Growth hormone 1).
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PLACENTA;
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible], Golos T.G.;
RT   "[illegible] growth hormone/chorionic somatomammotropin-related
RT   complementary deoxyribonucleic acids differentially expressed during
RT   pregnancy in the rhesus monkey placenta.";
RL   Endocrinology 13[illegible]:1744-17[illegible](199[illegible]).
CC   -!- FUNCTION: GROWTH HORMONE PLAYS AN IMPORTANT ROLE IN GROWTH
CC       CONTROL.
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
DR   EMBL; [illegible]; AAA[illegible].1; -.
DR   HSSP; [illegible]; [illegible].
DR   InterPro; IPR001400; Somatotropin.
DR   Pfam; PF00103; hormone; 1.
DR   PRINTS; PR00836; SOMATOTROPIN.
DR   PROSITE; PS00[illegible]; SOMATOTROPIN_2; UNKNOWN_1.
KW   [illegible]; Hormone; Signal.
FT   SIGNAL       1     2[illegible]
FT   CHAIN        2[illegible]   217       SOMATOTROPIN 1.
FT   DISULFID     7[illegible]   1[illegible]       BY SIMILARITY.
FT   DISULFID     20[illegible]   21[illegible]       BY SIMILARITY.
SQ   SEQUENCE   217 AA;  24942 MW;  FF5AA891[illegible]F2BC[illegible];

  Query Match             84.3%;  Score [illegible];  DB [illegible];  Length 217;
  Best Local Similarity   82.2%;  Pred. No. [illegible];
  Matches  [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy          [illegible] FIPLSRLFDNAMLRAHRLH[illegible]AFDTYQEFEEAYIPKEQKYSFLQNP[illegible]SLCFSESIPTP [illegible]2
Db         28 FSVIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPTP 87

Qy         [illegible] SNREETQQKSNLELLRISLLLIQSWLEPVQ [illegible]
Db         [illegible] SNREETQQKSNLELLRISLLLIQSWLEPVQ [illegible]


RESULT [illegible]
[illegible]
ID   [illegible]    PRELIMINARY;      PRT;   2[illegible] AA.
AC   [illegible];
DT   [illegible]-NOV-1999 (TrEMBLrel. [illegible], Created)
DT   [illegible]-NOV-1999 (TrEMBLrel. [illegible], Last sequence update)
DT   [illegible]-MAR-[illegible] (TrEMBLrel. [illegible], Last annotation update)
DE   [illegible] somatomammotropin [illegible] ([illegible] somatomammotropin [illegible]
DE   [illegible].
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible], [illegible] Y., Smith [illegible], [illegible] S., [illegible] A., [illegible];
RA   [illegible];
RT   "The human growth hormone [illegible] sequence [illegible] and
RT   [illegible].";
RL   Genomics 4:479-497(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible], Phillips J.A. III.;
```

```
RT   "[illegible]
RT   [illegible]";
RL   Science [illegible].
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PLACENTA;
RA   Strausberg R.;
RL   Submitted (JAN-[illegible]) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; [illegible]; AAA[illegible].
DR   EMBL; [illegible]; AAH[illegible].
DR   HSSP; [illegible]; [illegible].
DR   InterPro; [illegible].
DR   Pfam; PF[illegible].
DR   PRINTS; PR[illegible]; SOMATOTROPIN.
DR   PROSITE; PS00[illegible]; SOMATOTROPIN_1; 1.
DR   PROSITE; PS00[illegible]; SOMATOTROPIN_2; 1.
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible];

  Query Match             [illegible];  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches  [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy          [illegible] LSRLFDNAML[illegible]
Db         2[illegible] IVPLSRLFDHAMLQAHRAHQLAIDTYQEFEETYIPKDQKYSFLHDSQTSFCFSD[illegible]

Qy         [illegible] SNREETQQKSNLELLRISLLLIQSWLEPVQ [illegible]
Db         [illegible] NMEETQQKSNLELLRISLLLIESWLEPVR [illegible]


RESULT [illegible]
Q8WNI[illegible]
ID   Q8WNI[illegible]    PRELIMINARY;      PRT;   [illegible] AA.
AC   Q8WNI[illegible];
DT   01-MAR-2002 (TrEMBLrel. 2[illegible], Created)
DT   01-MAR-2002 (TrEMBLrel. 2[illegible], Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 2[illegible], Last annotation update)
DE   Growth hormone.
GN   GH V.
OS   Ateles geoffroyi (Black-handed spider monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Atelinae; Ateles.
OX   NCBI_TaxID=95[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RA   [illegible] A., [illegible], [illegible];
RT   "Independent duplication [illegible] of the growth hormone [illegible] in three
RT   Anthropoidean lineages.";
RL   Submitted (APR-[illegible]) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF[illegible]; AAL[illegible].1; -.
DR   InterPro; IPR001400; Somatotropin.
DR   Pfam; PF00103; hormone; 1.
DR   PRINTS; PR00836; SOMATOTROPIN.
DR   PROSITE; PS00[illegible]; SOMATOTROPIN_1; UNKNOWN_1.
DR   PROSITE; PS00[illegible]; SOMATOTROPIN_2; UNKNOWN_1.
SQ   SEQUENCE   21[illegible] AA;  [illegible] MW;  [illegible];

  Query Match             [illegible];  Score [illegible];  DB [illegible];  Length 2[illegible];
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches  [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy          [illegible] FIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFS[illegible]
Db         2[illegible] FPPLSRLFDNAMLRAHRL[illegible]

Qy         [illegible] SNREETQQKSNLELLRISLLLIQSWLEPVQ [illegible]
Db         8[illegible] PENREEVIAKSNLELLRISLLLIQSWLEPVQ [illegible]


RESULT [illegible]
[illegible]
```

```
ID   [illegible]     PRELIMINARY;     PRT;   202 AA.
AC   [illegible];
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Placental growth hormone 20kDa isoform precursor.
GN   [illegible].
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=FULL-TERM PLACENTA;
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible], [illegible], [illegible], [illegible];
RA   Carlsson L.M.S., Carlsson B.;
RT   "Cloning of two novel growth hormone transcripts expressed in human
RT   placenta.";
RL   J. Clin. Endocrinol. Metab. [illegible](1998).
DR   EMBL; AF[illegible]; AAB71828.1; -.
DR   HSSP; [illegible]; 1A22.
DR   InterPro; IPR001400; Somatotropin.
DR   Pfam; PF00103; hormone; 1.
DR   PRINTS; [illegible]; SOMATOTROPIN.
DR   PROSITE; PS00266; SOMATOTROPIN_1; 1.
DR   PROSITE; PS00338; SOMATOTROPIN_2; 1.
KW   Signal.
FT   SIGNAL        1     26       POTENTIAL.
SQ   SEQUENCE   202 AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             71.[illegible]%;  Score [illegible];  DB 4;  Length 202;
  Best Local Similarity   76.9%;  Pred. No. 6.7e-29;
  Matches   70;  Conservative    [illegible];  Mismatches    [illegible];  Indels   15;  Gaps    1;

Qy          2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPT 61
Db         27 FPTIPLSRLFDNAMLRARRLYQLAYDTYQEF                NPQTSLCFSESIPT  [illegible]

Qy         62 PSNREETQQKSNLELLRISLLLIQSWLEPVQ 92
Db         [illegible] PSNRVKTQQKSNLELLRISLLLIQSWLEPVQ 102


RESULT [illegible]
[illegible]
ID   [illegible]     PRELIMINARY;     PRT;   216 AA.
AC   [illegible];
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Growth hormone precursor.
OS   Spalax leucodon ehrenbergi (Ehrenberg's mole rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Spalacinae;
OC   Nannospalax.
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Iliopoulos A., Nevo E., Wallis M.;
RT   "Cloning and characterisation of the gene encoding mole rat (Spalax
RT   ehrenbergi) growth hormone.";
RL   J. Mol. Endocrinol. 22:29-36(1999).
DR   EMBL; AJ[illegible]; CAA[illegible]6.1; -.
DR   HSSP; P01241; 1BST.
DR   InterPro; IPR001400; Somatotropin.
DR   Pfam; PF00103; hormone; 1.
DR   PRINTS; [illegible]; SOMATOTROPIN.
DR   PROSITE; PS00266; SOMATOTROPIN_1; 1.
DR   PROSITE; PS00338; SOMATOTROPIN_2; 1.
KW   Signal.
FT   SIGNAL        1     26       POTENTIAL.
FT   CHAIN        27    216       GROWTH HORMONE.
SQ   SEQUENCE   216 AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             [illegible]%;  Score [illegible];  DB 11;  Length 216;
  Best Local Similarity   [illegible]%;  Pred. No. [illegible];
  Matches   [illegible];  Conservative    [illegible];  Mismatches   17;  Indels    [illegible];  Gaps    [illegible];

Qy          2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPT [illegible]
Db         27 FPAMPLSNLFANAVLRAQHLHQLAADTYKEFERAYIPEGQRYSIQNAQAAFCFSETIPA [illegible]

Qy         62 PSNREETQQKSNLELLRISLLLIQSWLEPVQ [illegible]
Db         [illegible] PSGKEEAQQRSDMELLRFSLLLIQSWLGPVQ [illegible]


RESULT 10
Q9R2C3
ID   Q9R2C3     PRELIMINARY;     PRT;   216 AA.
AC   Q9R2C3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   Growth hormone.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Nguyen T.N.K., Liebhaber S.A.;
RT   "Mouse Growth Hormone Locus: Nucleotide Sequence and Phylogenetic
RT   Analyses.";
RL   Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U34362; AAC99088.1; -.
DR   HSSP; P01241; 1BST.
DR   InterPro; IPR001400; Somatotropin.
DR   Pfam; PF00103; hormone; 1.
DR   PRINTS; PR00836; SOMATOTROPIN.
DR   PROSITE; PS00266; SOMATOTROPIN_1; 1.
DR   PROSITE; PS00338; SOMATOTROPIN_2; 1.
SQ   SEQUENCE   216 AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             [illegible]%;  Score [illegible];  DB 11;  Length 216;
  Best Local Similarity   [illegible]%;  Pred. No. [illegible];
  Matches   58;  Conservative   14;  Mismatches   18;  Indels    [illegible];  Gaps    [illegible];

Qy          2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPT [illegible]
Db         27 FPAMPLSSLFSNAVLRAQHLHQLAADTYKEFERAYIPEGQRYSIQNAQAAFCFSETIPA [illegible]

Qy         62 PSNREETQQKSNLELLRISLLLIQSWLEPVQ [illegible]
Db         [illegible] PTGKEEAQQRTDMELLRFSLLLIQSWLGPVQ [illegible]


RESULT 11
[illegible]
ID   [illegible]     PRELIMINARY;     PRT;   [illegible] AA.
AC   [illegible];
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Placental growth hormone (Fragment).
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; [illegible]
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PLACENTA;
RX   MEDLINE=[illegible]; PubMed=[illegible];
```



Run on: February 5, 2003, 12:42:26 ; Search time 5.5167 Seconds
(without alignments)
205.219 Million cell updates/sec

Title: US-09-423-100-2
Perfect score: 470
Sequence: 1 MSPQTLSREINAMLRAAn.....NLRLwvSEIL;SWrtV, 4

Scoring table: BL_SUM62
Xgap 10.0, Xgext 0.5

Searched: 2625/4 seqs, 241.2922 residues

Total number of hits satisfying chosen parameters: 2624

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/PCT_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*
5: /cgn2_6/ptodata/1/iaa/backfiles2.pep:*

Query No. 1: The number of results predicted by chance to have a
score greater than or equal to the score is the E-value printed.
Results determined by analysis of the query versus all sequences in the

SUMMARIES

	Result No.	Score	Query Match	Length	DB ID	Description
	1	85	18.1	102	1	US-09-284-028A-1 Sequence 1, Appli
	2	84	17.9	141	1	US-09-284-028A-2 Sequence 2, Appli
	3	84	17.9	144	2	US-08-[illegible] Sequence 3, Appl
	4	84	17.9	144	2	US-08-[illegible] Sequence 4, Appl
	5	84	17.9	217	1	US-08-[illegible] Sequence 5, Appl
	6	84	17.9	217	1	US-08-[illegible] Sequence 6, Appl
	7	84	17.9	217	1	US-08-[illegible] Sequence 7, Appl
	8	84	17.9	217	1	US-08-284-028A-1 Sequence 1, Appli
	9	84	17.9	218	4	US-09-424-[illegible] Sequence 25, Appl
	10	84	17.9	274	2	US-08-284-028A-2 Sequence 71, Appl
	11	84	17.9	300	2	US-08-284-028A-3 Sequence 73, Appl
	12	84	17.9	341	1	US-[illegible] Sequence 1, Appli
	13	84	17.9	217	1	US-[illegible] Sequence 4, Appli
	14	84	17.9	217	1	US-[illegible] Sequence 5, Appl
	15	84	17.9	217	2	US-08-[illegible] Sequence 1, Appli
	16	83	17.7	141	4	US-08-[illegible] Sequence 18, Appl
	17	83	17.7	141	4	US-08-[illegible] Sequence 19, Appl
	18	83	17.7	141	4	US-[illegible] Sequence 20, Appl
	19	83	17.7	172	1	US-[illegible] Sequence 1, Appli
	20	83	17.7	172	1	US-[illegible] Sequence 3, Appl
	21	83	17.7	172	1	US-[illegible] Sequence 4, Appl
	22	83	17.7	172	2	US-[illegible] Sequence 4, Appli
	23	83	17.7	1[illegible]	4	Patent No. 5424[illegible]
	24	[illegible]	[illegible]	148	2	US-[illegible] Sequence 5, Appl
	25	[illegible]	[illegible]	148	2	US-[illegible] Sequence 5, Appli
	26	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	27	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	28	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	29	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	30	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	31	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	32	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	33	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	34	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	35	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	36	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	37	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	38	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	39	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	40	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	41	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	42	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	43	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	44	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	45	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

ALIGNMENTS

RESULT 1
US-09-423-100-2
; Sequence 1, Application US/09423100
; Patent No. 6468[illegible]
; GENERAL INFORMATION:
; APPLICANT: DePuy, Bernard A.
; APPLICANT: Bostrom, Bertil T.
; APPLICANT: Sterett, Peter B.
; TITLE OF INVENTION: DNA Encoding Expressible Human Bone Morphogenetic
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: [illegible]
; STREET: [illegible]
; CITY: South San Francisco
; STATE: [illegible]
; COUNTRY: USA
; ZIP: [illegible]
; COMPUTER READABLE FORM:
; MEDIUM TYPE: [illegible]
; COMPUTER: [illegible]
; OPERATING SYSTEM: [illegible]
; SOFTWARE: PatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/[illegible]
; FILING DATE: [illegible]
; CLASSIFICATION: [illegible]
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: [illegible]
; FILING DATE: [illegible]
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: [illegible]
; FILING DATE: [illegible]
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: [illegible]
; FILING DATE: [illegible]
; ATTORNEY/AGENT INFORMATION:
; NAME: [illegible], Reed A.
; REGISTRATION NUMBER: [illegible]
; REFERENCE/DOCKET NUMBER: [illegible]
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: [illegible]
; TELEFAX: [illegible]
; TELEX: [illegible]
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: [illegible]
; TYPE: amino [illegible]

; Type: Linear
; US-08-xxx-xxx-1

Query Match 98.9%; Score 465; DB 1; Length 192;
Best Local Similarity 98.9%; Pred. No. 1.1e-51;
Matches 91; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MAYPLRSALANAFLVHNLHAMHHDLFDFQKYLQWGYSSNRNQVFYGNPQISLSFSLSETF 60
Db 1 MAYPLRSALANAFLVHNLHAMHHDLFDFQKYLQWGYSSNRNQVFYGNPQISLSFSLSETF 60

Qy 61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQ 92
Db 61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQ 92

RESULT 2
US-09-284-878-5
; Sequence 5, Application US/09284878
; Patent No. 6342375
; GENERAL INFORMATION:
; APPLICANT: Wittmann, Martha Guerrero
; APPLICANT: Soldana, Hugo Barrera
; APPLICANT: Salvado, Jose Maria Vinader
; TITLE OF INVENTION: Genetically Modified Methylotrophic P. pastoris Yeast for the
; TITLE OF INVENTION: Production and Secretion of the Human Growth Hormone
; FILE REFERENCE: 1829.0010000
; CURRENT APPLICATION NUMBER: US/09/234,878
; CURRENT FILING DATE: 1999-01-21
; PRIOR APPLICATION NUMBER: ES 9702254
; PRIOR FILING DATE: 1997-10-24
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-284-878-5

Query Match 97.9%; Score 460; DB 4; Length 191;
Best Local Similarity 98.9%; Pred. No. 4.6e-51;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPT 61
Db 2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPT 61

Qy 62 PSNREETQQKSNLELLRISLLLIQSWLEPVQ 92
Db 62 PSNREETQQKSNLELLRISLLLIQSWLEPVQ 92

RESULT 3
US-08-xxx-621-4
; Sequence 4, Application US/08383621
; Patent No. 5917072
; GENERAL INFORMATION:
; APPLICANT: Daley, Michael J.
; APPLICANT: Hackbarth, Brian L.
; APPLICANT: Cady, Susan M.
; APPLICANT: Shieh, Hong-Ming
; APPLICANT: Bohlen, Peter
; APPLICANT: Seddon, Andrew P.
; TITLE OF INVENTION: Stabilization of Somatotropins and Other
; TITLE OF INVENTION: Proteins by Modification of Cysteine Residues
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: American Cyanamid Company
; STREET: One Cyanamid Plaza
; CITY: Wayne
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 07470-8426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/766,112
; FILING DATE: 25-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Tsevdos, Estelle J.
; REGISTRATION NUMBER: 31,145
; REFERENCE/DOCKET NUMBER: 31,278-F1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 203-821-2756
; TELEFAX: 203-821-2971
; TELEX: 203-710-474-4059
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 194 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-383-621-4

Query Match 97.9%; Score 460; DB 2; Length 194;
Best Local Similarity 98.9%; Pred. No. 4.7e-51;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPT 61
Db 4 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPT 63

Qy 62 PSNREETQQKSNLELLRISLLLIQSWLEPVQ 92
Db 64 PSNREETQQKSNLELLRISLLLIQSWLEPVQ 94

RESULT 4
US-08-459-906-4
; Sequence 4, Application US/08459906
; Patent No. 6010999
; GENERAL INFORMATION:
; APPLICANT: Daley, Michael J.
; APPLICANT: Hackbarth, Brian L.
; APPLICANT: Cady, Susan M.
; APPLICANT: Shieh, Hong-Ming
; APPLICANT: Bohlen, Peter
; APPLICANT: Seddon, Andrew P.
; TITLE OF INVENTION: Stabilization of Somatotropins and Other
; TITLE OF INVENTION: Proteins by Modification of Cysteine Residues
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: American Cyanamid Company
; STREET: One Cyanamid Plaza
; CITY: Wayne
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 07470-8426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,906
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Darryl L.

```
RESULT 12
US-08-423-100-2
; Sequence 2, Application US/08423100
; Patent No. 5759760
; GENERAL INFORMATION:
;  APPLICANT: Newkirk, Christopher S.
;  APPLICANT: Halenbeck, Phillipe A.
;  APPLICANT: Nelson, Eileen Hamilton
;  APPLICANT: Clark, Samuel A.
;  APPLICANT: Thirpen, Alice E.
;  APPLICANT: Christian
;  APPLICANT: Krause, Fred
;  APPLICANT: McGarry, Dennis
;  TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM
;  TITLE OF INVENTION: SECRETORY CELL LINES
;  NUMBER OF SEQUENCES: 79
;  CORRESPONDENCE ADDRESS:
;    ADDRESSEE: Arnold, White & Durkee
;    STREET: P.O. Box 4433
;    CITY: Houston
;    STATE: Texas
;    COUNTRY: USA
;    ZIP: 77210
;  COMPUTER READABLE FORM:
;    MEDIUM TYPE: Floppy disk
;    COMPUTER: IBM PC compatible
;    OPERATING SYSTEM: PC-DOS/MS-DOS
;    SOFTWARE: Patentin Release #1.0, Version #1.30
;  CURRENT APPLICATION DATA:
;    APPLICATION NUMBER: US/08/423,100
;    FILING DATE: Concurrently Herewith
;    CLASSIFICATION: 435
;  PRIOR APPLICATION DATA:
;    APPLICATION NUMBER: US 60/010,427
;    FILING DATE: 15-JAN-1996
;  PRIOR APPLICATION DATA:
;    APPLICATION NUMBER: 08/471,779
;    FILING DATE: 06-JUN-1995
;  ATTORNEY/AGENT INFORMATION:
;    NAME: Alexander, Steven L.
;    REGISTRATION NUMBER: 37,842
;    REFERENCE/DOCKET NUMBER: 0150.5113
;  TELECOMMUNICATION INFORMATION:
;    TELEPHONE: 512/418-3000
;    TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 79:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 98 amino acids
    TYPE: amino acid
    STRANDEDNESS:
    TOPOLOGY: linear
US-08-423-100-2

  Query Match                 96.9%;  Score 454;  DB 1;  Length 191;
  Best Local Similarity   97.0%;  Pred. No. 2.7e-50;
  Matches   96;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 EPTLESKFLLNAMLPAHRLHQLAFDTYQEFEAAYIPKDQKYSFLQNPQTSLCFSESIPTP 61
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2 EPTLESKFLLNAMLPAHRLHQLAFDTYQEFEAAYIPKDQKYSFLQNPQTSLCFSESIPTP 61

Qy      62 PSNKEETQQKSNLELLRISLLLIQSWLEPVQ 93
           |||||||||||||||||||||||||||||||
Db      62 PSNKEETQQKSNLELLRISLLLIQSWLEPVQ 93


RESULT 13
US-08-423-100-2
; Sequence 2, Application US/08423100
; Patent No. 5759760
; GENERAL INFORMATION:
;  APPLICANT: Newkirk, Christopher S.
;  APPLICANT: Halenbeck, Phillipe A.
;  APPLICANT: Nelson, Eileen Hamilton
;  APPLICANT: Clark, Samuel A.
;  APPLICANT: Thirpen, Alice E.
;  APPLICANT: Christian
;  APPLICANT: Krause, Fred
;  APPLICANT: McGarry, Dennis
;  TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM
;  TITLE OF INVENTION: SECRETORY CELL LINES
;  NUMBER OF SEQUENCES: 79
;  CORRESPONDENCE ADDRESS:
;    ADDRESSEE: Arnold, White & Durkee
;    STREET: P.O. Box 4433
;    CITY: Houston
;    STATE: Texas
;    COUNTRY: USA
;    ZIP: 77210
;  COMPUTER READABLE FORM:
;    MEDIUM TYPE: Floppy disk
;    COMPUTER: IBM PC compatible
;    OPERATING SYSTEM: PC-DOS/MS-DOS
;    SOFTWARE: Patentin Release #1.0, Version #1.30
;  CURRENT APPLICATION DATA:
;    APPLICATION NUMBER: US/08/423,100
;    FILING DATE: Concurrently Herewith
;    CLASSIFICATION: 435
;  PRIOR APPLICATION DATA:
;    APPLICATION NUMBER: US 60/010,427
;    FILING DATE: 15-JAN-1996
;  PRIOR APPLICATION DATA:
;    APPLICATION NUMBER: 08/471,779
;    FILING DATE: 06-JUN-1995
;  ATTORNEY/AGENT INFORMATION:
;    NAME: Alexander, Steven L.
;    REGISTRATION NUMBER: 37,842
;    REFERENCE/DOCKET NUMBER: 0150.5113
;  TELECOMMUNICATION INFORMATION:
;    TELEPHONE: 512/418-3000
;    TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 79:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 98 amino acids
    TYPE: amino acid
    STRANDEDNESS:
    TOPOLOGY: linear
US-08-423-100-2

  Query Match                 96.9%;  Score 454;  DB 1;  Length 191;
  Best Local Similarity   97.0%;  Pred. No. 2.7e-50;
  Matches   96;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 EPTLESKFLLNAMLPAHRLHQLAFDTYQEFEAAYIPKDQKYSFLQNPQTSLCFSESIPTP 61
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2 EPTLESKFLLNAMLPAHRLHQLAFDTYQEFEAAYIPKDQKYSFLQNPQTSLCFSESIPTP 61

Qy      62 PSNKEETQQKSNLELLRISLLLIQSWLEPVQ 93
           |||||||||||||||||||||||||||||||
Db      62 PSNKEETQQKSNLELLRISLLLIQSWLEPVQ 93


RESULT 14
US-08-423-100-4
; Sequence 4, Application US/08423100
; Patent No. 5759760
; GENERAL INFORMATION:
;  APPLICANT: BOSEN, EI AL.
;  TITLE OF INVENTION: Human Growth Hormone
;  NUMBER OF SEQUENCES: 7
;  CORRESPONDENCE ADDRESS:
;    ADDRESSEE: Amgen Inc.
;    STREET: 1840 DeHavilland Drive
;    CITY: Thousand Oaks
;    STATE: CA
;    COUNTRY: USA
;    ZIP: 91320-1789
;  COMPUTER READABLE FORM:
;    MEDIUM TYPE: 3.5 inch diskette
;    COMPUTER: IBM PC/AT
;    OPERATING SYSTEM: MS-DOS
;    SOFTWARE: WordPerfect 5.1
;  CURRENT APPLICATION DATA:
;    APPLICATION NUMBER: US/08/423,100
;    FILING DATE: Herewith
;    CLASSIFICATION: 435
;  PRIOR APPLICATION DATA:
;    APPLICATION NUMBER:
;    FILING DATE:
;  ATTORNEY/AGENT INFORMATION:
;    NAME: Perman, Steven D.
;    REGISTRATION NUMBER: 32,008
;    REFERENCE/DOCKET NUMBER: A-273-CIP
;  TELECOMMUNICATION INFORMATION:
;    TELEPHONE: 201 994 1700
;    TELEFAX: 201 994 1744
INFORMATION FOR SEQ ID NO: 4:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 217 amino acids
    TYPE: amino acid
    STRANDEDNESS:
    TOPOLOGY: linear
US-08-423-100-4
```

Comparison protein search, using sw model

Run on: February 5, 2003, 12:44:31 ; Search time 0.24878 Seconds
(without alignments)
247.409 Million cell updates/sec

Title: US-09-423-100-2
Perfect score: 470
Sequence: 1 MEPIIPLSKLEDNAMIAAE.....NLELLPLNLAL_WETV_ 92

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 129505 seqs, 22169297 residues

Total number of hits satisfying chosen parameters: 129505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubaa/US07_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubaa/US07_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubaa/US08_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubaa/US08_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubaa/US09A_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubaa/US09B_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubaa/US60_PUBCOMB.pep:*

First, No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	470	100.0	92	9	US-10-054-873-2	Sequence 2, Appli
2	470	100.0	150	9	US-10-054-873-7	Sequence 7, Appli
3	466	97.4	217	10	US-09-423-100-2	Sequence 2, Appli
4	466	97.4	245	10	US-09-290-[illegible]	Sequence 66, Appl
5	451	96.0	217	10	US-09-854-088-2	Sequence 2, Appli
6	450	95.7	217	9	US-10-864-[illegible]	Sequence 16, Appl
7	447	95.1	217	10	US-09-853-[illegible]	Sequence 4, Appli
8	446	94.9	217	10	US-09-850-887-3	Sequence 3, Appli
9	265	56.4	49	9	US-10-054-873-1	Sequence 1, Appli
10	260	55.3	107	9	US-10-051-873-6	Sequence 6, Appli
11	259.5	55.2	466	10	US-09-887-500A-2	Sequence 2, Appli
12	112	23.8	54	10	US-09-876-478-14	Sequence 14, Appl
13	105	22.3	198	9	US-10-[illegible]	Sequence 32, App
14	104	22.1	199	9	US-10-[illegible]	Sequence 33, Appl
15	104	22.1	199	9	US-10-140-[illegible]	Sequence 26, Appl
16	126	26.8	199	9	US-10-140-[illegible]	Sequence 27, Appl
17	124	26.4	199	9	US-10-140-[illegible]	Sequence 31, Appl
18	123	26.2	199	9	US-10-[illegible]	Sequence 22, Appl
19	122	26.0	199	9	US-10-140-[illegible]	Sequence 21, Appl
20	121.5	25.9	62	10	US-[illegible]	Sequence 1, Appl
21	121.5	25.9	254	10	US-[illegible]	Sequence 1, Appl
22	119	25.3	199	9	US-10-[illegible]	Sequence 23, Appl
23	118	25.1	199	9	US-10-[illegible]	Sequence 24, Appl
24	118	25.1	199	9	US-10-140-[illegible]	Sequence 25, Appl
25	117	24.9	197	9	US-10-140-[illegible]	Sequence 29, Appl
26	116	24.7	197	9	US-10-140-[illegible]	Sequence 30, Appl
27	116	24.7	199	9	US-10-140-[illegible]	Sequence 28, Appl
28	116	24.7	199	9	US-10-140-[illegible]	Sequence 32, Appl
29	116	24.7	199	9	US-10-140-[illegible]	Sequence 33, Appl
30	116	24.7	227	9	US-10-140-[illegible]	Sequence 5, Appl
31	116	24.7	227	9	US-10-140-[illegible]	Sequence 4, Appl
32	116	24.7	227	10	US-[illegible]	Sequence 1, Appl
33	115	24.5	199	9	US-10-140-[illegible]	Sequence 3, Appl
34	115	24.5	199	9	US-10-140-[illegible]	Sequence 1, Appl
35	114	24.3	54	10	US-09-[illegible]	Sequence 1, Appl
36	114	24.3	54	10	US-09-[illegible]	Sequence 2, Appl
37	114	24.3	199	9	US-10-140-[illegible]	Sequence 12, Appl
38	112	23.8	199	9	US-10-[illegible]	Sequence 13, Appl
39	111	23.6	199	9	US-10-[illegible]	Sequence 14, Appl
40	106	22.6	28	10	US-09-[illegible]	Sequence 1, Appl
41	106	22.6	28	10	US-09-[illegible]	Sequence 2, Appl
42	106	22.6	54	10	US-09-[illegible]	Sequence 3, Appl
43	101	21.5	198	9	US-10-[illegible]	Sequence 34, Appl
44	101	21.5	197	9	US-10-[illegible]	Sequence 35, Appl
45	99	21.1	199	9	US-10-[illegible]	Sequence 36, Appl

ALIGNMENTS

RESULT 1
US-10-054-873-2
; Sequence 2, Application US/10054873
; Publication No. US20030124681A1
; GENERAL INFORMATION:
; Applicant: Lehrach, Hans
; TITLE OF INVENTION: Libraries for Full-Length Containing an Optimized Universal Phage Protein-Like Sequence
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: [illegible]
; STREET: San Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/054,873
; FILING DATE: 22-Jan-2002
; CLASSIFICATION: Unknown
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: [illegible]
; FILING DATE: 23-Mar-[illegible]
; APPLICATION NUMBER: [illegible]
; FILING DATE: 11-[illegible]
; ATTORNEY/AGENT INFORMATION:
; NAME: Mercer, Frank
; REGISTRATION NUMBER: 43,343
; REFERENCE/DOCKET NUMBER: [illegible]
; TELECOMMUNICATION INFORMATION:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 92 amino acids
; TYPE: amino acid
; STRANDEDNESS: Unknown
; TOPOLOGY: Linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

; Sequence 7, Application US/10/423,871
; Patent No. US20020197712A1
; GENERAL INFORMATION:
; APPLICANT: Guo, Zhong Fu
; TITLE OF INVENTION: Numeric Protein Classification
; Using Hydrophobic Amphotropic Like Sequence
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/423,871
; FILING DATE: 22-Jan-2002
; CLASSIFICATION: Unknown
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/287,615
; FILING DATE: 01-MAY-2001
; APPLICATION NUMBER: US 09/12x,xxx
; FILING DATE: 15-APR-2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Mycroft, Frank J
; REGISTRATION NUMBER: 46,946
; REFERENCE/DOCKET NUMBER: 020167-000130US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-326-2400
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 150 amino acids
; TYPE: amino acid
; STRANDEDNESS: Unknown
; TOPOLOGY: Linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:

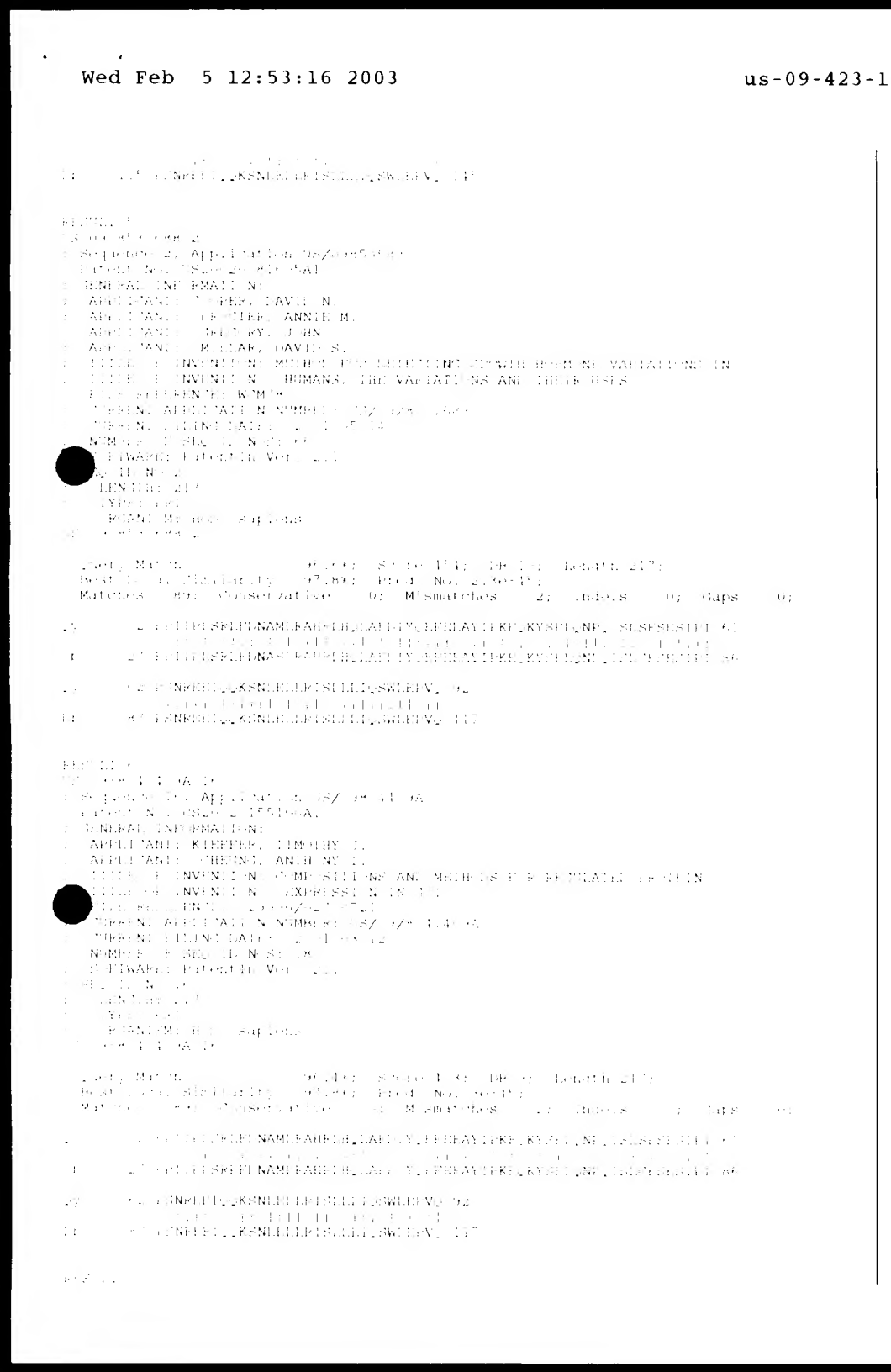
; Sequence 8, Application US/10/423,871
; Patent No. US20020197712A1
; GENERAL INFORMATION:
; APPLICANT: Guo, Zhong Fu
; TITLE OF INVENTION: Numeric Protein Classification
; Using Hydrophobic Amphotropic Like Sequence
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/423,871
; FILING DATE: 22-Jan-2002
; CLASSIFICATION: Unknown
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/287,615
; FILING DATE: 01-MAY-2001
; APPLICATION NUMBER: US 09/12x,xxx
; FILING DATE: 15-APR-2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Mycroft, Frank J
; REGISTRATION NUMBER: 46,946
; REFERENCE/DOCKET NUMBER: 020167-000130US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-326-2400
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 150 amino acids
; TYPE: amino acid
; STRANDEDNESS: Unknown
; TOPOLOGY: Linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:

; Sequence 9, Application US/10/423,871
; Patent No. US20020197712A1
; GENERAL INFORMATION:
; APPLICANT: Guo, Zhong Fu
; TITLE OF INVENTION: Numeric Protein Classification
; Using Hydrophobic Amphotropic Like Sequence
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/423,871
; FILING DATE: 22-Jan-2002
; CLASSIFICATION: Unknown
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/287,615
; FILING DATE: 01-MAY-2001
; APPLICATION NUMBER: US 09/12x,xxx
; FILING DATE: 15-APR-2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Mycroft, Frank J
; REGISTRATION NUMBER: 46,946
; REFERENCE/DOCKET NUMBER: 020167-000130US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-326-2400
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 150 amino acids
; TYPE: amino acid
; STRANDEDNESS: Unknown
; TOPOLOGY: Linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 9:

; Sequence 10, Application US/10/423,871
; Patent No. US20020197712A1
; GENERAL INFORMATION:
; APPLICANT: Guo, Zhong Fu
; TITLE OF INVENTION: Numeric Protein Classification
; Using Hydrophobic Amphotropic Like Sequence
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/423,871
; FILING DATE: 22-Jan-2002
; CLASSIFICATION: Unknown
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/287,615
; FILING DATE: 01-MAY-2001
; APPLICATION NUMBER: US 09/12x,xxx
; FILING DATE: 15-APR-2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Mycroft, Frank J
; REGISTRATION NUMBER: 46,946
; REFERENCE/DOCKET NUMBER: 020167-000130US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-326-2400
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 150 amino acids
; TYPE: amino acid
; STRANDEDNESS: Unknown
; TOPOLOGY: Linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:

; Sequence 11, Application US/10/423,871
; Patent No. US20020197712A1
; GENERAL INFORMATION:
; APPLICANT: Guo, Zhong Fu
; TITLE OF INVENTION: Numeric Protein Classification
; Using Hydrophobic Amphotropic Like Sequence
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/423,871
; FILING DATE: 22-Jan-2002
; CLASSIFICATION: Unknown
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/287,615
; FILING DATE: 01-MAY-2001
; APPLICATION NUMBER: US 09/12x,xxx
; FILING DATE: 15-APR-2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Mycroft, Frank J
; REGISTRATION NUMBER: 46,946
; REFERENCE/DOCKET NUMBER: 020167-000130US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-326-2400
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 150 amino acids
; TYPE: amino acid
; STRANDEDNESS: Unknown
; TOPOLOGY: Linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 11:

; Sequence 12, Application US/10/423,871
; Patent No. US20020197712A1
; GENERAL INFORMATION:
; APPLICANT: Guo, Zhong Fu
; TITLE OF INVENTION: Numeric Protein Classification
; Using Hydrophobic Amphotropic Like Sequence
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/423,871
; FILING DATE: 22-Jan-2002
; CLASSIFICATION: Unknown
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/287,615
; FILING DATE: 01-MAY-2001
; APPLICATION NUMBER: US 09/12x,xxx
; FILING DATE: 15-APR-2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Mycroft, Frank J
; REGISTRATION NUMBER: 46,946
; REFERENCE/DOCKET NUMBER: 020167-000130US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-326-2400
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 150 amino acids
; TYPE: amino acid
; STRANDEDNESS: Unknown
; TOPOLOGY: Linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:



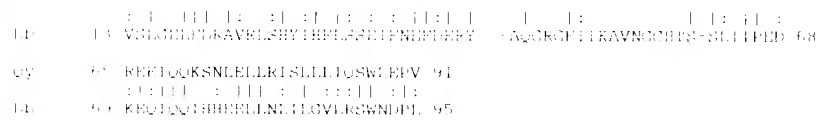
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 5 ;

Query Match 84.3%; Score 396; DB 10; Length 217;
Best Local Similarity 94.6%; Pred. No. 1.2e-88;
Matches 77; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

; PATENT NO. US20020164712A1 ;
; GENERAL INFORMATION:
; APPLICANT: Gao, Zhaoli P.
; TITLE OF INVENTION: Cytokine Protein Variation and
; Intramolecular Chaperone-Like Sequence
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon LLP
; STREET: 633 West Fifth Street, 47th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94071-3804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/054,077
; FILING DATE: 22-Jan-2002
; CLASSIFICATION: Unknown
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/423,100
; FILING DATE: 11-DEC-2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Frank J
; REGISTRATION NUMBER: 44,949
; REFERENCE/DOCKET NUMBER: 203.7 [illegible]US
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 79 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:

Query Match 55.9%; Score 262; DB 9; Length 1[illegible];
Best Local Similarity 100.0%; Pred. No. 2.8e-24;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 12
US/09/887,569A-2
; Sequence 2, Application US/09/887,569A
; Patent No. US20020100078A1
; GENERAL INFORMATION:
; APPLICANT: McCuaig, Marion K.
; APPLICANT: Habib, Hamid R.
; TITLE OF INVENTION: Expression of Somatotropin in Prunt Seeds
; FILE REFERENCE: 9369.DP3
; CURRENT APPLICATION NUMBER: US/09/887,569A
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: US 60/212,[illegible]
; PRIOR FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 99
; TYPE: PRT



Search Completed: February 5, 2003, 12:48:47
Job Time : 8.24373 secs

Migration protein search, using sw model

Run on: February 5, 2003, 12:42:21 ; Search time 45.1998 Seconds
(without alignments)
526.115 Million cell updates/sec

[redacted] US-09-423-100-4
Molecule type: 463
Query: [redacted]

Scoring: [redacted]

Searching: [redacted]

Total number of hits satisfying chosen parameters: [redacted]

Minimum DB seq length: [redacted]
Maximum DB seq length: [redacted]

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/gcgp[redacted]
2: [redacted]
3: [redacted]
4: [redacted]
5: [redacted]
6: [redacted]
7: [redacted]
8: [redacted]
9: [redacted]
10: [redacted]
11: [redacted]
12: [redacted]
13: [redacted]
14: [redacted]
15: [redacted]
16: [redacted]
17: [redacted]
18: [redacted]
19: [redacted]
20: [redacted]
21: [redacted]
22: [redacted]
23: [redacted]
24: [redacted]
25: [redacted]
26: [redacted]
27: [redacted]
28: [redacted]
29: [redacted]
30: [redacted]
31: [redacted]
32: [redacted]
33: [redacted]
34: [redacted]
35: [redacted]
36: [redacted]
37: [redacted]
38: [redacted]
39: [redacted]
40: [redacted]
41: [redacted]
42: [redacted]
43: [redacted]
44: [redacted]
45: [redacted]

[redacted]

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	[redacted]					Sequence 2 human
2	[redacted]					Human insulin pr
3	[redacted]					Human insulin pr
4	[redacted]					Human preproinsu
5	[redacted]					Human preproinsu
6	[redacted]					Human preproinsu
7	[redacted]					Sequence 1 from
8	[redacted]					Sequence 1 from
9	[redacted]					Sequence 1 from
10	[redacted]					Human preproinsu
11	[redacted]					Human preproinsu
12	[redacted]					Human preproinsu
13	[redacted]					Human preproinsu
14	[redacted]					Human preproinsu
15	[redacted]					Human preproinsu
16	[redacted]					Human preproinsu
17	[redacted]					Human preproinsu
18	[redacted]					Human preproinsu
19	[redacted]					Human preproinsu
20	[redacted]					Human preproinsu
21	[redacted]					Human preproinsu
22	[redacted]					Human preproinsu
23	[redacted]					Human preproinsu
24	[redacted]					Human preproinsu
25	[redacted]					Human preproinsu
26	[redacted]					Human preproinsu
27	[redacted]					Human preproinsu
28	[redacted]					Human preproinsu
29	[redacted]					Human preproinsu
30	[redacted]					Human preproinsu
31	[redacted]					Human preproinsu
32	[redacted]					Human preproinsu
33	[redacted]					Human preproinsu
34	[redacted]					Human preproinsu
35	[redacted]					Human preproinsu
36	[redacted]					Human preproinsu
37	[redacted]					Human preproinsu
38	[redacted]					Human preproinsu
39	[redacted]					Human preproinsu
40	[redacted]					Human preproinsu
41	[redacted]					Human preproinsu
42	[redacted]					Human preproinsu
43	[redacted]					Human preproinsu
44	[redacted]					Human preproinsu
45	[redacted]					Human preproinsu

ALIGNMENTS

[redacted]

```
ID   ... 1998 bp.
XX
AC   ...;
XX
DT   ... (first entry)
XX
DE   Human insulin.
XX
KW   Insulin transformation gene expression; fusion; tumour cell;
KW   cytotoxic A chain; cytokine; immunosuppressive.
XX
OS   Homo sapiens.
XX
PN   WO9...-...-A1.
XX
PD   ... 1999.
XX
PF   ... MAR ...-WO-...
XX
PR   ... MAR 1998; 98WO-...
XX
PA   (DENI ) DENIVA PHARM BV.
XX
PI   ...;
XX
DR   WPI; ...-...../...
XX
PT   New chimeric proteins containing insulin, or its precursor fragments, and
PT   particularly for the production of human insulin
XX
PS   Claim 1; ... pp; ... English.
XX
CC   This sequence represents a human insulin precursor comprising insulin
CC   A and B chains separated by a 34 residue peptide sequence. This novel
CC   precursor can be ... of choice of proteins which ... cells.
CC   Systems ... N terminal fragment of human growth hormone (hGH) and a
```

```
ID   ... 1998 bp.
XX
AC   ...;
XX
DT   ... (first entry)
XX
DE   Human insulin.
XX
KW   Insulin transformation gene expression; fusion; tumour cell;
KW   cytotoxic A chain; cytokine; immunosuppressive.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1...2..
FT                   /*tag= a
FT                   /product= "insulin."
XX
CN   ... A.
XX
CC   ... ...
XX
CC   ... ...
XX
CC   ... ...
XX
CC   ... ...
XX
PA   (MEST/) MESTREC S.
PA   (ZELV ) ZELVA PHARM & ...
XX
```

```
DT   ... (first entry)
XX
DE   Human insulin precursor, SEQ ID ...
XX
KW   Insulin precursor; ... hormone; ... gene; ... disease; ...
KW   ... cDNA; ... immunogenic; ... ...; ... ...; ...
KW   production; ...
XX
OS   Homo sapiens.
XX
PN   WO9...-...-A1.
XX
PD   07-OCT-1999.
XX
PF   ... MAR ...-WO-...
XX
PR   ... MAR 1998; 98WO-...
XX
PA   (DENI ) ... BV.
XX
IN   Gan Z.
XX
DR   WPI; ...-...../...
XX
PT   New chimeric proteins containing insulin, or its precursor fragments, and
PT   particularly for the production of human insulin
XX
PS   Claim 1; ... pp; English.
XX
CC   This sequence represents a human insulin precursor comprising insulin
CC   A and B chains separated by a 34 residue peptide sequence. This novel
CC   precursor can be ... of choice of proteins which ... cells.
CC   Systems ... N terminal fragment of human growth hormone (hGH) and a
```

cleavable peptide linker (AAY42857). The hGH portion of the chimeric protein acts as an intramolecular chaperone (IMC) for the insulin precursor, enabling it to fold correctly. The cleavable peptide linker has a C-terminal Arg residue which enables the hGH portion of the 'fusion' protein to be removed after folding has taken place. Production of recombinant human insulin via an hGH-proinsulin chimeric protein can be human insulin with correctly linked cysteine residues with fewer necessary procedural steps, and hence results in a higher yield than the other IMC systems not using protease sequences to facilitate proteolytic degradation by a host organism host, but also promote the folding of the fused insulin precursor, facilitate the solubility of the fusion protein and decrease the intermolecular interactions among the fusion proteins, thus preventing or inhibiting the undesired folding of the fused insulin precursor. Additionally useful high concentrations of mercaptan or the use of high pH (greater than about pH 10.5) sulphite lysis and related purification steps can thus be eliminated, along with the use of high concentrations of mercaptan or the use of hydrophobic absorbent resins.

XX

SQ Sequence ** AA;

Query Match Score ; Length ;
Best Local Similarity ; Pred. No. ;
Matches ; Conservative ; Mismatches ; Indels ; Gaps ;

Qy

Db

RESULT 1
AAM43...
ID AAM43... standard protein ** AA.
XX
AC AAM43...;
XX
DT ... (first entry)
XX
DE Human preproinsulin protein sequence SEQ ID NO:...
XX
KW ...
KW ...
KW ...
KW ...
KW ...
KW ...
KW ...
KW ...
KW ...
KW ...
KW ...
XX
OS ...
XX
FH ...
FT ...
XX
SQ Sequence ** AA;

Query Match ...
Best Local Similarity ...
Matches ...

Qy

Db

RESULT 2
AAM43...
ID AAM43... standard protein ** AA.
XX
AC AAM43...;
XX
DT ... (first entry)
XX
DE Human preproinsulin...
XX
KW ...
KW ...
KW ...
XX
OS ...
XX
FH ...
FT ...
XX
SQ Sequence ** AA;

Query Match ...
Best Local Similarity ...
Matches ...

Qy

Db

```
RT   Prevention and treatment of cartilage disorders including cartilage damage by injury, a
RT   particular to cartilaginous disorders, by contacting cartilage with
RT   insulin-like growth factor analog with altered affinity for IGF binding
RT   proteins.
XX
RN   [1]
RP   1-85 Score 8.1 Ex: Jepp: English.
XX
PR   The present invention relates to a method for treating cartilage
PR   disorders. The method comprises contacting cartilage with an active moiet
PR   y, such as insulin like growth factor (IGF-1) analog with a binding affinity
PR   preference for IGF binding protein-3 (IGFBP-3) over IGFBP-1, an IGF-1
PR   analog with a binding affinity preference for IGFBP-1 over IGFBP-3, or a
PR   IGFBP displacer peptide that prevents the interaction of IGF with an
PR   IGFBP and does not bind to human IGF receptor. The method is useful for
PR   treating cartilage disorders (CD), including degenerative CD, articular
PR   CD such as rheumatoid arthritis and osteoarthritis. The present sequence
PR   is a human proinsulin, which was used to illustrate the invention.
XX
SQ   Sequence  86 AA;

  Query Match             Score 45;  DB 2;  Length 86;
  Best Local Similarity   Prob. No. 1.3e-1;
  Matches  20;  Conservative  0;  Mismatches  4;  Indels  0;  Gaps  0;

Qy      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQ 70
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQ 71

Qy     71 SLQKPCTIVEQCCFRSCDLALLENYCN 98
          ||||||||||||||||||||||||||||
Db     72 SLQKPCTIVEQCCFRSCDLALLENYCN 99


RESULT 2
AAP40217
ID   AAP40217 standard; protein; 87 AA.
XX
AC   AAP40217;
XX
DT   12-DEC-2002  (first entry)
XX
DE   Sequence of the 2N N-terminal AAs of proinsulin.
XX
KW   Hormone; diabetes; cartilage disorder; phase 0 variant.
XX
OS   Homo sapiens.
XX
RN   [1]
RP   1-87
RX   DOI; 10.1016/S0140-6736(02)11280-6.
XX
RA   Chen Y.C., Lin C.Y., Chen H.L., Wu C.Y., Chang C.F.;
XX
RT   "Expression of trypsin-resistant cationic peptides in rice plants
RT   confers enhanced resistance to sheath blight disease caused by
RT   Rhizoctonia solani in the field";
XX
RL   Ann. Appl. Biol. 142:303-312(2003).
XX
DR   WPI; 2003-072464/07.
DR   N-PSDB; AAN14179.
XX
PT   Protecting bacteria from phage infection, by transforming with a
PT   cloning vector containing segment with restriction and modification
PT   activity.
XX
PS   Example; Fig 1a; 24pp; English.
XX
CC   Plasmid pIn-alpha-1 was digested with EcoRI and BamHI. Insertion of a
CC   1.2 kb fragment encoding the 2N-terminal amino acids of proinsulin in
CC   the construction of plrp24, the inventors claim a method for
CC   protecting bacteria from phage infection by transforming with
CC   cloning vector containing segment with restriction and modification
CC   activity. Produn. of plasmid pIn-alpha-1 or plrp2 which uses pIn-alpha
CC   produn. of plasmid pRR29 which uses a synthetic gene coding for the
CC   42 N-terminal AAs of proinsulin (see AAN14179).
XX
SQ   Sequence  87 AA;

  Query Match             Score 45;  DB 2;  Length 87;
  Best Local Similarity   Prob. No. 1.3e-1;
  Matches  20;  Conservative  0;  Mismatches  4;  Indels  0;  Gaps  0;

Qy      1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQ 70
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQ 71

Qy     71 SLQKPCTIVEQCCFRSCDLALLENYCN 98
          ||||||||||||||||||||||||||||
Db     72 SLQKPCTIVEQCCFRSCDLALLENYCN 99
```

```
Db      [illegible] FVN[illegible]HL[illegible]VEALYLV[illegible]GERGFFYTPK[illegible]RREAE[illegible]V[illegible]VEL[illegible]GAGS[illegible]LALE[illegible]

Qy      [illegible]KR[illegible]VR[illegible]ISI[illegible]Y[illegible]LENY[illegible]N [illegible]

Db      [illegible]KR[illegible]VR[illegible]ISI[illegible]Y[illegible]LENY[illegible]N [illegible]

RESULT [illegible]
AAP[illegible]
ID   AAP[illegible] standard; Protein; 87 AA.
XX
AC   AAP[illegible]
XX
DT   [illegible] (first entry)
XX
DE   Sequence of the [illegible] N-terminal AAs of proinsulin.
XX
KW   Selectable vector; autonomously replicating vector;
KW   expression vector.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Region          1..31
FT                   /label= A chain
FT   Region          32..[illegible]
FT                   /label= B chain
FT   Region          [illegible]..87
FT                   /label= A chain
XX
PN   [illegible] A.
XX
PD   [illegible]
XX
PF   04-MAR-1985;  85EP-0[illegible]
XX
PR   [illegible]-MAR-1984;  84US-[illegible]
PR   [illegible]  84DN-[illegible]
XX
PA   ([illegible])
XX
PI   [illegible]
XX
DR   WPI; 1985-224[illegible]/37.
DR   N-PSDB; AAN[illegible]
XX
PT   New recombinant DNA expression vector [illegible]
PT   replication and co-transcription generating polycistronic mRNA
XX
PS   Example; Fig 14; [illegible]pp; English.
XX
CC   The inventors claim a process for preparing selectable and
CC   autonomously replicating recombinant DNA expression vectors which
CC   comprise 1) a transcriptional and translational activating sequence
CC   which is in the reading frame of a nucleotide sequence which codes
CC   for a peptide or polypeptide; 2) a translational stop signal; 3) a
CC   translational start signal which is in the reading frame of a
CC   [illegible] sequence that codes for a functional polypeptide;
CC   and 4) an additional translational stop signal. The peptide or
CC   polypeptide coding sequence codes for [illegible] AAs, esp. AAP[illegible]
CC   The functional polypeptide is esp. growth hormone, human insulin,
CC   interferon and human tissue plasminogen activator.
CC   (Updated on 16-AUG-2002 to add missing OS field.)
XX
SQ   Sequence [illegible] AA;

  Query Match [illegible] Score [illegible] DB [illegible] Length [illegible]
  Best Local Similarity [illegible] Pred. No. [illegible]
  Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy      [illegible] FVN[illegible]HL[illegible]VEALYLV[illegible]GERGFFYTPK[illegible]RREAE[illegible]V[illegible]VEL[illegible]GAGS[illegible]LALE[illegible]
```

```
Db      [illegible] FVN[illegible]HL[illegible]VEALY[illegible]

Qy      [illegible]KR[illegible]VR[illegible]ISI[illegible]Y[illegible]LENY[illegible]N [illegible]

Db      [illegible]KR[illegible]VR[illegible]ISI[illegible]Y[illegible]LENY[illegible]N [illegible]

RESULT [illegible]
AAP[illegible]
ID   AAP[illegible] standard; protein; [illegible] AA.
XX
AC   AAP[illegible]
XX
DT   [illegible] (first entry)
DT   [illegible] (latest entry)
XX
DE   Synthetic proinsulin.
XX
KW   Proinsulin; [illegible]
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Region          [illegible]
FT                   /label= B chain
FT   Region          [illegible]
FT                   /label= C chain
FT   Region          [illegible]
FT                   /label= A chain
XX
PN   [illegible] A.
XX
PD   [illegible]
XX
PF   04-MAR-1985;  [illegible]
XX
PR   05-MAR-1984;  [illegible]
PR   [illegible]-1984;  [illegible]
PR   [illegible]-JAN-1985;  [illegible]
XX
PA   ([illegible])
XX
PI   Bailey BM, [illegible]
XX
DR   WPI; 1985-[illegible]/4[illegible]
DR   N-PSDB; AAN[illegible]
XX
PT   New [illegible] DNA expression [illegible]
PT   useful in producing [illegible] transformed [illegible]
XX
PS   Disclosure; Fig. 14; [illegible]pp; English.
XX
CC   The synthetic proinsulin gene is expressed in a new selectable and
CC   autonomously replicating recombinant DNA expression vector
CC   comprising a runaway replicon and a transcriptional and
CC   translational activating sequence in the reading frame of the
CC   proinsulin [illegible]
CC   stop signal. [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   simplified.
CC   (Updated on 1[illegible])
XX
SQ   Sequence [illegible] AA;

  Query Match [illegible]
  Best Local Similarity [illegible] Pred. No. [illegible]
  Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]
```

XX
AC A [redacted];
XX
DE Proinsulin protein sequence.
XX
KW Human; proinsulin; vector; pBR[illegible]; pTINS; [illegible]; pBR[illegible]AI proinsulin;
KW insulin analogue; type II; type II diabetes.
XX
OS Synthetic.
XX
PN WO9[illegible]A.
XX
PD 18-FEB-1993.
XX
PF 31-JUL-1992; 92WO-US[illegible].
XX
PR 01-AUG-1991; 91US-[illegible].
PR 30-JUL-1992; 92US-[illegible].
XX
PA (PFIZ) PFIZER INC.
PA (SCIO-) SCIOS INC.
XX
PI Ando R.D.; Lucas C.F.
XX
DR WPI; [illegible].
DR N-PSDB; AA[illegible].
XX
PT New hepato-selective and peripheral-selective human insulin
PT analogues and their corresp. DNA, for treatment of type I and
PT type II diabetes.
XX
PS Disclosure; Fig 2; 54pp; English.
XX
CC This sequence represents human proinsulin and was decoded from
CC the sequences given in AA[illegible]. The cDNA fragment coding for
CC proinsulin was inserted into plasmid vector pBR[illegible] and digested with
CC KpnI and HindIII. This resulted in the formation of the vector pTINS.
CC A fragment encoding amino acids 1-29 of GAI (see AA[illegible]) was inserted
CC into pTINS to give a plasmid which encodes a DNA sequence which
CC for amino acids 1-29 of GAI, an 8-amino-acid linker sequence and human
CC proinsulin. This plasmid, pBR[illegible]AI proinsulin, could be used in the
CC formation of insulin analogues which may be used in the treatment of
CC types I and II diabetes.
XX
SQ Sequence 87 AA;

  Query Match [illegible]; Score [illegible]; DB [illegible]; Length 87;
  Best Local Similarity [illegible]; Pred. No. [illegible]e-43;
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy [illegible]
Db [illegible]

RESULT 2
AAB[illegible]
ID AAB[illegible] standard; protein; 87 AA.
XX
AC AAB[illegible];
XX
DT [illegible] (first entry)
XX
DE [illegible]
XX
KW Human; proinsulin; vector; pBR[illegible]; pTINS; GAI; pBR[illegible]AI proinsulin;
KW insulin analogue; type II; type II diabetes.
XX
OS Synthetic.
XX
PN WO9[illegible]A.
XX
PD 18-FEB-1993.
XX
PF 31-JUL-1992; 92WO-US[illegible].
XX
PR 01-AUG-1991; 91US-[illegible].
PR 30-JUL-1992; 92US-[illegible].
XX
PA (PFIZ) PFIZER INC.
PA (SCIO-) SCIOS INC.
XX
PI Ando R.D.; Lucas C.F.
XX
DR WPI; [illegible].
DR N-PSDB; AAB[illegible].
XX
PT New hepato-selective and peripheral-selective human insulin
PT analogues and their corresp. DNA, for treatment of type I and
PT type II diabetes.
XX
PS Disclosure; Fig 2; 54pp; English.
XX
CC This sequence represents human proinsulin and was decoded from
CC the sequences given in AAB[illegible]. The cDNA fragment coding for
CC proinsulin was inserted into plasmid vector pBR[illegible] and digested with
CC KpnI and HindIII. This resulted in the formation of the vector pTINS.
CC A fragment encoding amino acids 1-29 of GAI (see AAB[illegible]) was inserted
CC into pTINS to give a plasmid which encodes a DNA sequence which
CC for amino acids 1-29 of GAI, an 8-amino-acid linker sequence and human
CC proinsulin. This plasmid, pBR[illegible]AI proinsulin, could be used in the
CC formation of insulin analogues which may be used in the treatment of
CC types I and II diabetes.
XX
SQ Sequence 87 AA;

  Query Match [illegible]; Score [illegible]; DB [illegible]; Length 87;
  Best Local Similarity [illegible]; Pred. No. [illegible]e-43;
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy [illegible]
Db [illegible]

RESULT 3
AAB[illegible]
ID AAB[illegible] standard; protein; 87 AA.
XX
AC AAB[illegible];
XX
DT 14-FEB-1995 (first entry)
XX
DE Modified human insulin precursor.
XX
KW Human; insulin precursor; synthesis.
XX
FH Key Location/Qualifiers
FT Peptide [illegible]

SW protein - protein search, using sw model

Run on: February 5, 2003, 12:42:21 ; Search time 14.1792 Seconds
(without alignments)
583.076 Million cell updates/sec

Title: US-09-423-100-4
Perfect score: 463
Sequence: 1 FVNQHLCGSHLVLALYLVCG.....IVEQCCTSICSLYQLENYCN 86

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	463	100.0	110	1	IPIG	insulin precursor
2	463	100.0	110	2	A42779	insulin precursor
3	456	98.5	110	2	B42779	insulin precursor
4	456	98.5	110	2	A96278	insulin precursor
5	414	89.6	110	1	INPB	insulin precursor
6	407	87.9	110	1	IPEG	insulin precursor
7	394	85.1	86	1	IPBG	insulin precursor
8	394	85.1	110	1	IPRC2	insulin 2 precurso
9	394	85.1	110	1	INMS2	insulin 2 precurso
10	392	84.7	108	2	A96845	insulin precursor
11	392	84.7	110	2	I48.6	insulin precursor
12	385	83.2	110	1	IPRI1	insulin 1 precurso
13	384	82.9	84	1	IPBO	insulin precursor
14	376.5	79.2	105	1	IPBC	insulin precursor
15	366	79.0	108	1	INMS1	insulin 1 precurso
16	363.5	78.5	108	2	S09278	insulin precursor
17	325.5	69.2	77	1	INSH	insulin precursor
18	314	67.8	110	1	IPGB	insulin precursor
19	277.5	59.9	109	1	IPRIO3	insulin precursor
20	274.5	59.3	105	2	I51221	insulin precursor
21	265.5	57.3	106	1	IPXC2	insulin II precurs
22	265.5	57.3	107	1	IPCH	insulin precursor
23	262.5	56.7	106	1	IPXC1	insulin I precurso
24	256.5	55.4	51	1	INWBP	insulin - sperm wh
25	256.5	55.4	51	1	INWBF	insulin - finback
26	256.5	55.4	51	1	INEL	insulin - elephant
27	256.5	55.4	81	1	IPDK	insulin precursor
28	256	55.3	96	2	B57282	epidermal growth f
29	254.5	55.0	51	1	INHY	insulin - hamster
30	251.5	54.3	51	1	INMSP	insulin - spalax
31	250.5	54.1	51	1	A57171	insulin - [illegible]
32	246.5	53.2	51	1	INWLS	insulin - [illegible]
33	246.5	53.2	51	1	INSI	insulin - [illegible]
34	246.5	53.2	51	1	INCMA	insulin - [illegible]
35	245.5	53.0	51	1	INCI	insulin - [illegible]
36	244.5	52.8	51	1	INMKS	insulin - [illegible]
37	239.5	51.7	51	2	[illegible]	insulin - [illegible]
38	234.5	50.6	51	1	INCB	insulin - [illegible]
39	231.5	50.0	51	1	INCS	insulin - [illegible]
40	227.5	49.1	51	1	INPJ	insulin - [illegible]
41	227.5	49.1	51	1	INIK	insulin - [illegible]
42	227.5	49.1	51	1	INOS	insulin - [illegible]
43	227.5	49.1	51	2	A57124	insulin - [illegible]
44	227.5	49.1	51	2	A57161	insulin - [illegible]
45	226	48.8	51	2	S53154	insulin - [illegible]

ALIGNMENTS

RESULT 1
IPBO
insulin precursor [validated] - human
N;Alternate names: preproinsulin
C;Species: Homo sapiens (man)
C;Date: 23-Oct-1981 #sequence_revision 23-Oct-1981 #text_change [illegible]
C;Accession: A95222; A96278; A95225; B96278; A95244; A92075; A95226; [illegible]
R;Bell, G.I.; Pictet, R.L.; Rutter, W.J.; Cordell, B.; Tischer [illegible] H.M.
Nature 284, 26-32, 1980
A;Title: Sequence of the human insulin gene.
A;Reference number: A95222; MUID:80127725; PMID:6243748
A;Accession: A95222
A;Molecule type: DNA
A;Residues: 1-110 <BEL>
A;Cross-references: GB:J00265; NID:g186; PIDN:AAA59172.1; [illegible]
R;Ullrich, A.; Bell, J.R.; Chen, E.; Herrera, R.; Petruzzelli, L.M.; Dull, T.J.; [illegible]
Science 209, 612-615, 1980
A;Title: Genetic variation in the human insulin gene.
A;Reference number: A96278; MUID:80305916; PMID:6248959
A;Accession: A96278
A;Molecule type: DNA
A;Residues: 1-110 <BEL>
A;Cross-references: GB:J00265; NID:g186; PIDN:AAA59172.1; [illegible]
R;Bell, G.I.; Swain, W.F.; Pictet, R.; O'Neill, P.; Bradford, B.M.; [illegible]
Nature 282, 525-527, 1979
A;Title: Nucleotide sequence of a cDNA clone encoding human preproinsulin
A;Reference number: A96278; MUID:80047177; PMID:[illegible]
A;Accession: A96278
A;Molecule type: mRNA
A;Residues: 1-110 <BEL>
A;Cross-references: GB:J00265; NID:g186; PIDN:AAA59172.1; [illegible]
R;Soares, L.; Seidel, L.V.; Gray, A.; Steiner, A.
Science 208, 57-59, 1980
A;Title: Nucleotide sequence of rat preproinsulin mRNA
A;Reference number: A94151; MUID:80215107; PMID:6247954
A;Accession: A94151
A;Molecule type: mRNA
A;Residues: 1-110 <SOA>
A;Cross-references: GB:[illegible]; NID:g[illegible]; PIDN:AAA[illegible]
R;Nicol, L.S.E.W.; Scott, J.
Nature 187, 483-485, 1960
A;Title: Amino-acid sequence [illegible]
A;Reference number: A95144
A;Accession: A95144
A;Molecule type: protein
A;Residues: 2-[illegible]; 31-[illegible]
R;Boyer, P.E.; Cha, S.; Peterson, [illegible]
J. Biol. Chem. 241, 1558-1564, 1966
A;Title: Studies on human proinsulin. II. [illegible]
A;Reference number: A[illegible]; MUID:[illegible]
A;Accession: A[illegible]
A;Molecule type: protein

A;Residues: 57-87 <OYE>
R;Ko, A.; Smyth, D.G.; Markussen, J.; Sundby, F.
Eur. J. Biochem. 20, 190-199, 1971
A;Title: Amino acid sequence of the C-peptide of human proinsulin.
A;Reference number: A91186; MUID:71257722; PMID:5569404
A;Accession: A91186
A;Molecule type: protein
A;Residues: 57-87 <KOA
R;Bennett, A.M.; Julier, C.; Bellanné, J.P.; Bellard, S.; Froguel, P.; Lathrop, M.; Bell
Nature Genet. 4, 405-410, 1993
A;Title: Susceptibility to insulin dependent diabetes mellitus maps to a 4.1 kb segment
A;Reference number: 158111; MUID:94364428; PMID:8358443
A;Accession: 158114
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-59,63-110 <RES>
A;Cross-references: GB:L15440; NID:g307071; PIDN:AAA59179.1; PID:g307072
R;Sieber, P.; Kamber, B.; Hartmann, A.; Jöhl, A.; Riniker, B.; Rittel, W.
Helv. Chim. Acta 57, 2617-2621, 1974
A;Title: Totalsynthese von Humaninsulin unter gezielter Bildung der Disulfidbindungen.
A;Reference number: A91626; MUID:75077277; PMID:4443293
A;Contents: annotation; synthesis
A;Note: disulfide-bonded human insulin was synthesized; the synthetic hormone was identi
A;Note: article in German with English abstract
R;Nalbandi, V.K.
Hoppe-Seyler's Z. Physiol. Chem. 354, 659-672, 1973
A;Title: The synthesis of C-peptide of human proinsulin.
A;Reference number: A91659; MUID:73040097; PMID:4804504
A;Contents: annotation; synthesis of residues 57-87
R;Geiger, R.; Jaeger, G.; Koenig, W.
Chem. Ber. 106, 2447-2452, 1973
A;Title: Synthesis of the complete sequence of human proinsulin C-peptide and its [Glu
A;Reference number: A90914
A;Contents: annotation; synthesis of residues 57-87
R;Kaufmann, J.E.; Irminger, J.C.; Halban, P.A.
Biochem. J. 310, 869-874, 1995
A;Title: Sequence requirements for proinsulin processing at the B-chain/C-peptide junct
A;Reference number: S58661; MUID:96014185; PMID:7575420
A;Contents: annotation; site-directed mutagenesis study of proteolytic processing
C;Genetics:
A;Gene: GDB:INS
A;Cross-reference: GDB:119849; OMIM:176730
A;Map position: 11p15.5-11p15.5
A;Introns: 63/1
C;Superfamily: insulin
C;Keywords: hormone; pancreas
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-54/Domain: insulin chain B #status experimental <BCH>
F;25-54,90-110/Product: insulin #status experimental <MAT>
F;57-87/Domain: connecting C-peptide #status experimental <CPEP>
F;90-110/Domain: insulin chain A #status experimental <ACH>
F;31-96,44-109,95-100/Disulfide bonds: #status experimental

Query Match 100.0%; Score 463; DB 1; Length 110;
Best Local Similarity 100.0%; Pred. No. 2.9e-43;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 86
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 25 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 110

RESULT 2
A42179
insulin precursor - chimpanzee
C;Species: Pan troglodytes (chimpanzee)
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 16-Jul-1999
C;Accession: A42179; S22058
R;Seino, S.; Bell, G.I.; Li, W.H.
Mol. Biol. Evol. 9, 193-203, 1992
A;Title: Sequences of primate insulin genes support the hypothesis of a slower rate
A;Reference number: A42179; MUID:92219054; PMID:1560757
A;Accession: A42179
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-110 <SEI>
A;Cross-reference: EMBL:X61992; NID:g22203; PIDN:CAA43405.1; PID:g22204
A;Note: sequence obtained from NCBI backbone (NCBIN:95185, NCBIP:95194)
C;Genetics:
A;Introns: 63/1
C;Superfamily: insulin

Query Match 100.0%; Score 463; DB 2; Length 110;
Best Local Similarity 100.0%; Pred. No. 2.9e-43;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 86
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 25 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 110

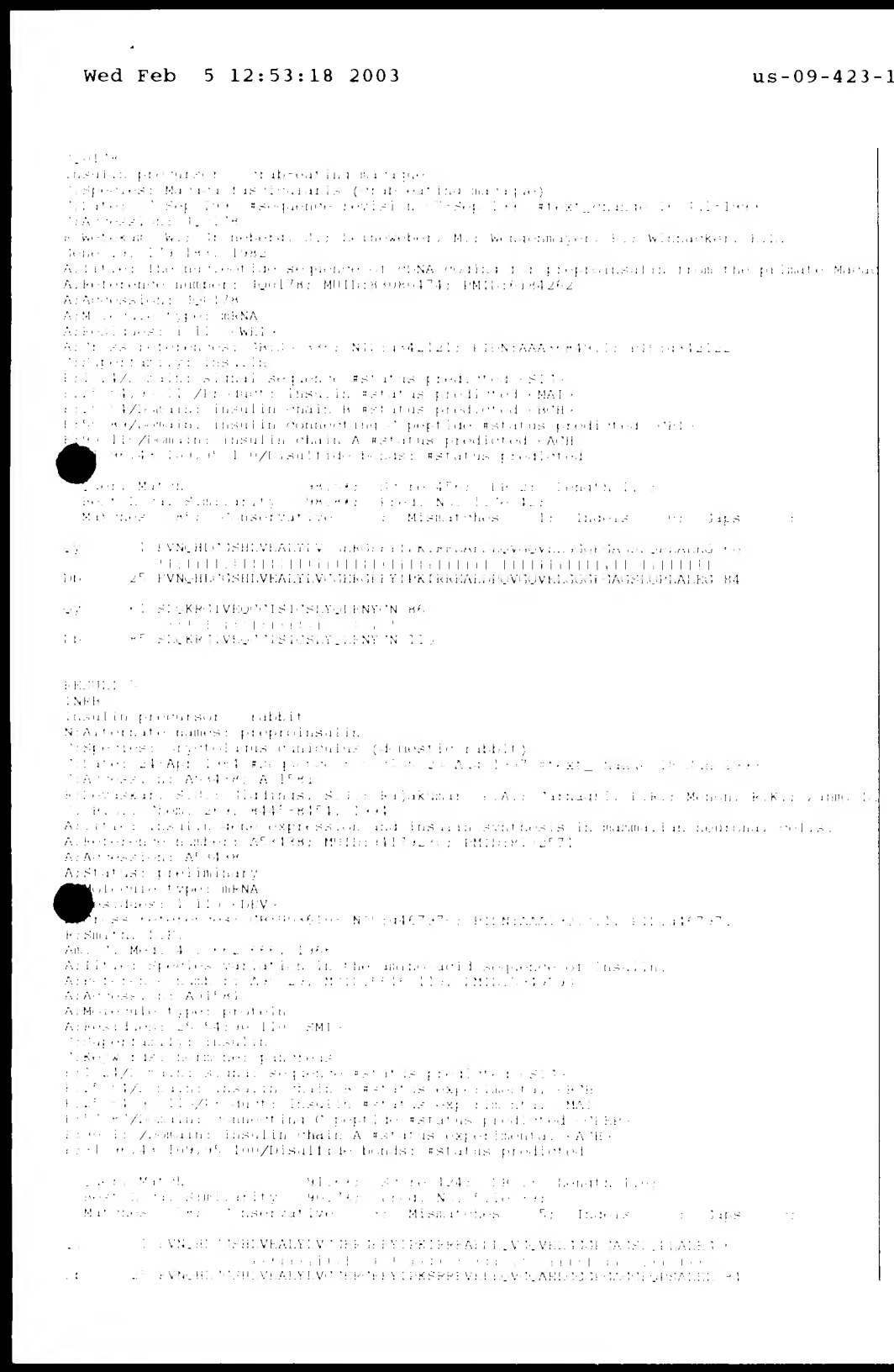
RESULT 3
B42179
insulin precursor - green monkey
C;Species: Cercopithecus aethiops (green monkey, grivet)
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 16-Jul-1999
C;Accession: B42179, A05282; S16494; S22058
R;Seino, S.; Bell, G.I.; Li, W.H.
Mol. Biol. Evol. 9, 193-203, 1992
A;Title: Sequences of primate insulin genes support the hypothesis of a slower rate
A;Reference number: A42179; MUID:92219054; PMID:1560757
A;Accession: B42179
A;Molecule type: DNA
A;Residues: 1-110 <SEI>
A;Cross-references: EMBL:X61992; NID:g22203; PIDN:CAA43405.1; PID:g22204
A;Note: sequence obtained from NCBI backbone (NCBIN:95185, NCBIP:95194)
R;Peterson, J.D.; Nehrlich, S.; Iyer, P.E.; Steiner, D.F.
J. Biol. Chem. 247, 4866-4871, 1972
A;Title: Determination of the amino acid sequence of the monkey, sheep, and Bos grunn
A;Reference number: A92111; MUID:72258016; PMID:4626469
A;Accession: A05282
A;Molecule type: protein
A;Residues: 57-87 <PEI>
C;Genetics:
A;Introns: 63/1
C;Superfamily: insulin
C;Keywords: hormone; pancreas
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-54/Domain: insulin chain B #status predicted <BCH>
F;25-54,90-110/Product: insulin #status predicted <MAT>
F;57-87/Domain: connecting peptide #status experimental <CPEP>
F;90-110/Domain: insulin chain A #status predicted <ACH>
F;31-96,44-109,95-100/Disulfide bonds: #status predicted

Query Match 99.4%; Score 460; DB 2; Length 110;
Best Local Similarity 98.8%; Pred. No. 1.7e-42;
Matches 85; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FVNQ[illegible] 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 25 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 110

Qy 61 SLQKR 86
|||||
Db 85 SLQKRGIVEQCCTSICSLYQLENYCN 110

RESULT 4



A;Accession: A04172
A;Molecule type: protein
A;Residues: 1-84 <BAT>
R;Sanger, F.; Kitai, R.
Nature 174, 556-565, 1954
A;Title: The structure of pig and sheep insulins.
A;Reference number: A90044
A;Accession: S10492
A;Molecule type: protein
A;Residues: 1-50;<1-51 <BRO>
R;Short, J.; Dahmaard, H.
Hoppe-Seyler's Z. Physiol. Chem. 20, 476-480, 1988
A;Title: Proinsulin heterogeneity in pigs.
A;Reference number: A60887; MUID:89062178; PMID:3184979
A;Accession: A00005
A;Molecule type: protein
A;Residues: 90-86;<1-2 <SNE>
A;Note: The A-chain is reported to be characteristic of a connecting peptide cortical hormone
A;Accession: A00005
A;Molecule type: protein
A;Residues: 33-5... <SNZ>
A;Cross-ref: L.F. Hodson, S.; Bodnikin, D.; Mercovia, D.
Adv. Protein Chem. 26, 279-362, 1972
A;Title: Insulin, the structure in the crystal and its reflection in chemistry and biolo
A;Reference number: A90017
A;Contents: annotation: X-ray crystallography, Lys mustrooms
C;Superfamily: insulin
C;Keywords: hormone; pancreas
F;1-50/Domain: insulin chain B #status experimental <INB>
F;52-81/Product: insulin #status experimental <INA>
F;52-56/Domain: connecting peptide #status experimental <CPE>
F;1-84/Domain: insulin chain A #status experimental <INA>
F;7...-11... 8...-9... 24/disulfide bonds: #status experimental

  Query Match 82.7%; Score 383; DB 1; Length 84;
  Best Local Similarity 89.0%; Pred. No. 1.1e-34;
  Matches 74; Conservative 1; Mismatches 9; Indels 2; Gaps 1;

Qy 1 FVN...
Db 1 FVN...

Qy ... 86
Db ... 84

RESULT 2
INMS1
Insulin precursor - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 21-Apr-1984 #sequence_revision 22-Apr-1995 #text_change 16-Jul-1999
C;Accession: A04172; A00005; A00017; A90044; A60887; A90017; A90018; A90019; A90020; A...
R;Steiner, D.; Yonnes, M.A.; White, J.W.; Peters, T.K.; Floyd, J.C., Jr.; Lo, T.; M...
Mol. Endocrinol. 1, 427-431, 1987
A;Title: Cloned and nucleotide sequence analysis of a supplementary bovine preproinsul...
A;Reference number: A43990; MUID:88208220; PMID:3453...
A;Accession: A10004
A;Molecule type: mRNA
A;Residues: 1-110 <DAA>
A;Cross-reference: GB:M51479; NID:g163478; PIDN:AAA...; GI:gi163479
A;Experimental source: fetal pancreas
R;Nolan, C.; Margoliash, E.; Peterson, J.D.; Steiner, D.F.
J. Biol. Chem. 246, 2780-2795, 1971
A;Title: The structure of bovine proinsulin.
A;Reference number: A92080; MUID:71166442; PMID:4928802
A;Accession: A92080
A;Molecule type: protein
A;Residues: 25-107 <NOL>
R;Oyer, P.E.; Horris, S.; Yen, P.E.; Lorris, S.; Peterson, J.D.; Steiner, D.F.
J. Biol. Chem. 246, 1641-1644, 1971
A;Title: Studies on rat insulin precursors. I. Purification and properties of two
A;Reference number: A92081; MUID:71166443; PMID:4928803

```
;Species: Mus musculus (house mouse)
;Date: 24-Apr-1984 #Sequence_revision 14-Jul-1994 #text_change 16-Jun-1999
;Accession: B26842; A48172; A01992; B60012
R;Wentworth, B.M.; Schaefer, I.M.; Villa-Komaroff, L.; Chirgwin, J.M.
J. Mol. Evol. 24, 305-312, 1986
A;Title: Characterization of the two nonallelic genes encoding mouse preproinsulin.
A;Reference number: A22665; MUID:87157763; PMID:3094005
A;Accession: B26842
A;Molecule type: DNA
A;Residues: 1-108 <WEN>
A;Cross-references: GB:X04724; NID:g52712; PIDN:CAA28434.1; PID:g52713
R;Soria, B.; Skelin, L.; Bonner-Weir, S.; Weir, G.C.
J. Mol. Endocrinol. 5, 61-67, 1990
A;Title: Molecular cloning and DNA sequence analysis of preproinsulin genes in the New M
A;Reference number: A48172; MUID:91072994; PMID:2407123
A;Accession: A48172
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-108 <SAW>
A;Cross-references: GB:M10242; NID:g190155; PIDN:AAA37392.1; PID:g190156
R;Brandenburg, D.; Gattner, H.G.; Kunz, P.; Moelhaupt, E.; Humbel, R.E.
Hoppe Seyler's Z. Physiol. Chem. 353, 451-458, 1972
A;Title: Amino acid sequence of the two insulins from mouse (Mus musculus).
A;Reference number: A01992; MUID:72184155; PMID:5063718
A;Accession: A01992
A;Molecule type: protein
A;Residues: 25-54;80 <BUE>
R;Glinde, E.; Nicholson, J.M.; Ransom, D.; Wellinder, M.S.
J. Chromat. 11, 9-21, 254, 1980
A;Title: Reverse-phase high performance liquid chromatographic analyses of insulin biosy
A;Reference number: A60012; MUID:80202078; PMID:7001585
A;Accession: B60012
A;Molecule type: protein
A;Residues: 57 <GLN>
;Superfamily: insulin
;Keywords: hormone; pancreas
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-54/Domain: insulin chain B #status experimental <MAT>
F;55-64,65-90/Product: insulin #status experimental <MAT>
F;57-64/Domain: connecting peptide #status experimental <CPP>
F;65-90/Domain: insulin chain A #status experimental <MAT>
F;31-91,43-109/disulfide bonds: #status predicted

  Query Match          70.0%;  Score 65;  DB 1;  Length 108;
  Best Local Similarity   81.4%;  Pred. No. 1e-12;
  Matches   57;  Conservative    4;  Mismatches    1;  Indels    2;  Gaps    1;

Qy       1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
           | |:|||:|||||:|||||||||||:||| || || |:||| | | || ||||
Db      25 FVKQHLCGPHLVEALYLVCGERGFFYTPKSRREVEDPQVEQLELGGSP--GDLQTLALEV 82

Qy      61 SLQKRGIVEQCCTSICSLYQLENYCN 86
           | :|||:|||||||||||:|||||
Db      83 ALQKRGIVDQCCTSICSLYQLENYCN 108


Search completed: February 5, 2003, 12:45:01
Job time : 5.1792 secs
```

Title: US-09-423-100-4
Perfect score: 463
Sequence: 1 PVNLHDLSGHLVLALVLV22.....[VL]TLLVLLGIENCPN 96

Searched: [illegible] seqs, [illegible] residues

Total number of hits satisfying chosen parameters: [illegible]

Minimum DB seq length:
Maximum DB seq length: 2[illegible]

Post-processing: Minimum Match [illegible]%
Maximum Match [illegible]%
Listing first [illegible] summaries

Database : SwissProt_40:*

Query No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

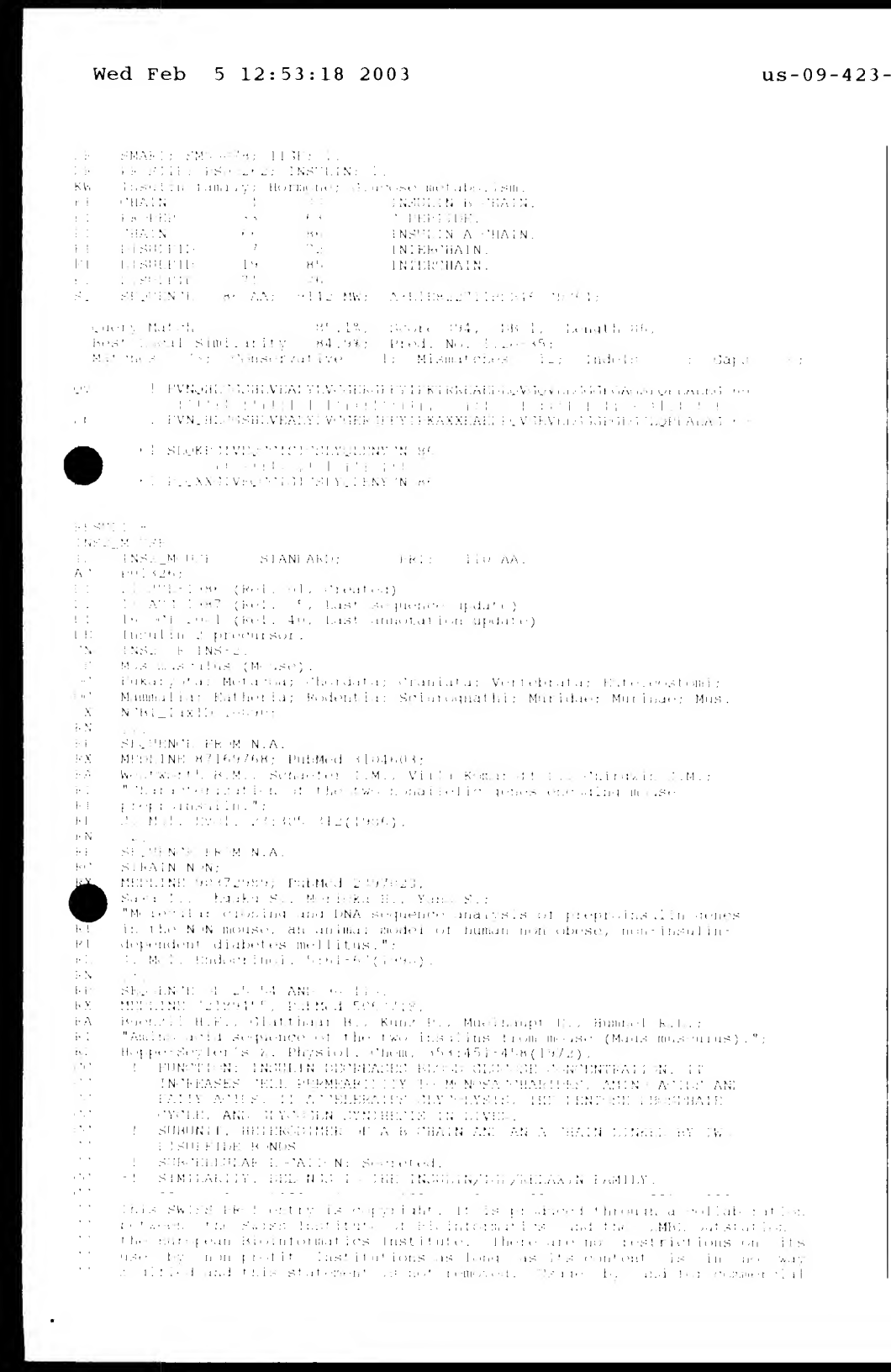
Result No.	Score	Match	Length	DB	ID	Description
1	463	100.0	96	1	INS_HUMAN	P01308 INSULIN PRECU
2	453	97.9	110	1	INS_PANTR	P49811 PANCREATIC PRE
3	450	97.2	110	1	INS_PERAE	P30407 PROINSULIN PE
4	436	94.3	110	1	INS_MACFA	P30406 MACAQUE INSUL
5	424	91.6	110	1	INS_RABIT	P01311 PREPROINSULIN
6	417	90.1	110	1	INS_CANFA	P01321 CANIS FAMILI
7	394	85.1	86	1	INS_HORSE	P01311 EQUUS CABALL
8	394	85.1	110	1	INS_MOUSE	P01325 MUS MUSCULU
9	394	85.1	110	1	INS2_RAT	P01323 RATTUS NORV
10	392	84.7	104	1	INS_ALLMI	P01314 ALLOMYRINA
11	392	84.7	11	1	INS_PHILO	P14175 PHILODINA
12	389	84.0	11	1	INS_RAT	P01322 INSULIN-LI
13	389	84.0	110	1	INS_LAMA	P01317 LOS LAMINS
14	386	83.4	110	1	INS_BOVIN	P01317 BOS TAURUS
15	386	83.4	110	1	INS1_MOUSE	P01325 MUS MUSCULU
16	384.5	83.0	110	1	INS_SHEEP	P01315 OVIS ARIES
17	384.5	83.0	110	1	INS_PIG	P01315 SUBUNIT OF
18	384	82.9	110	1	INS_MACAV	P01329 MACACA MULA
19	382.5	82.6	110	1	INS_CHICK	P01319 GALLUS GALL
20	382.5	82.6	110	1	INS2_XENLA	P12087 XENOPUS LAE
21	382.5	82.6	110	1	INS_ONCMY	P01322 ONCORHYNCHU
22	382.5	82.6	110	1	INS1_XENLA	P12087 XENOPUS LAE
23	381.5	82.4	51	1	INS_BALPH	P01312 BALAENOPTER
24	381.5	82.4	51	1	INS_ELEMA	P01318 ELEPHUS MAX
25	381.5	82.4	51	1	INS_AMIGU	P01324 AMIA CALVA
26	381.5	82.4	51	1	INS_ALLMA	P01314 ALLOMYRINA
27	381.5	82.4	51	1	INS_PAPHA	P01316 PAPIO HAMA
28	381.5	82.4	51	1	INS_CAMDR	P01320 CAMELUS DR
29	381.5	82.4	51	1	INS_VARNI	P01326 VARANUS NI
30	381.5	82.4	51	1	INS_EQUCA	P01317 EQUUS CABA
31	381.5	82.4	51	1	INS_CAIMA	P01320 CAIMAN CRO

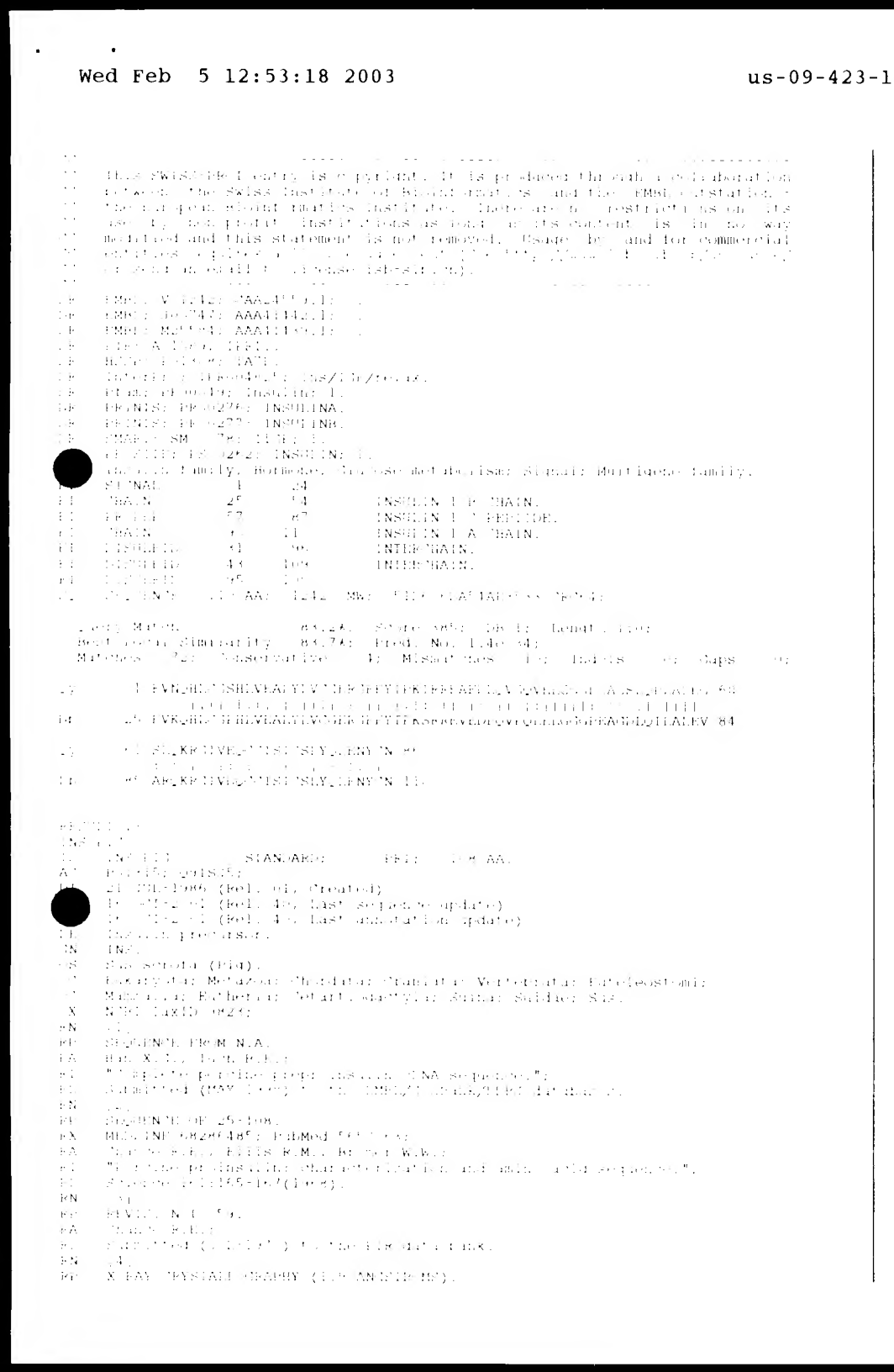
ALIGNMENTS

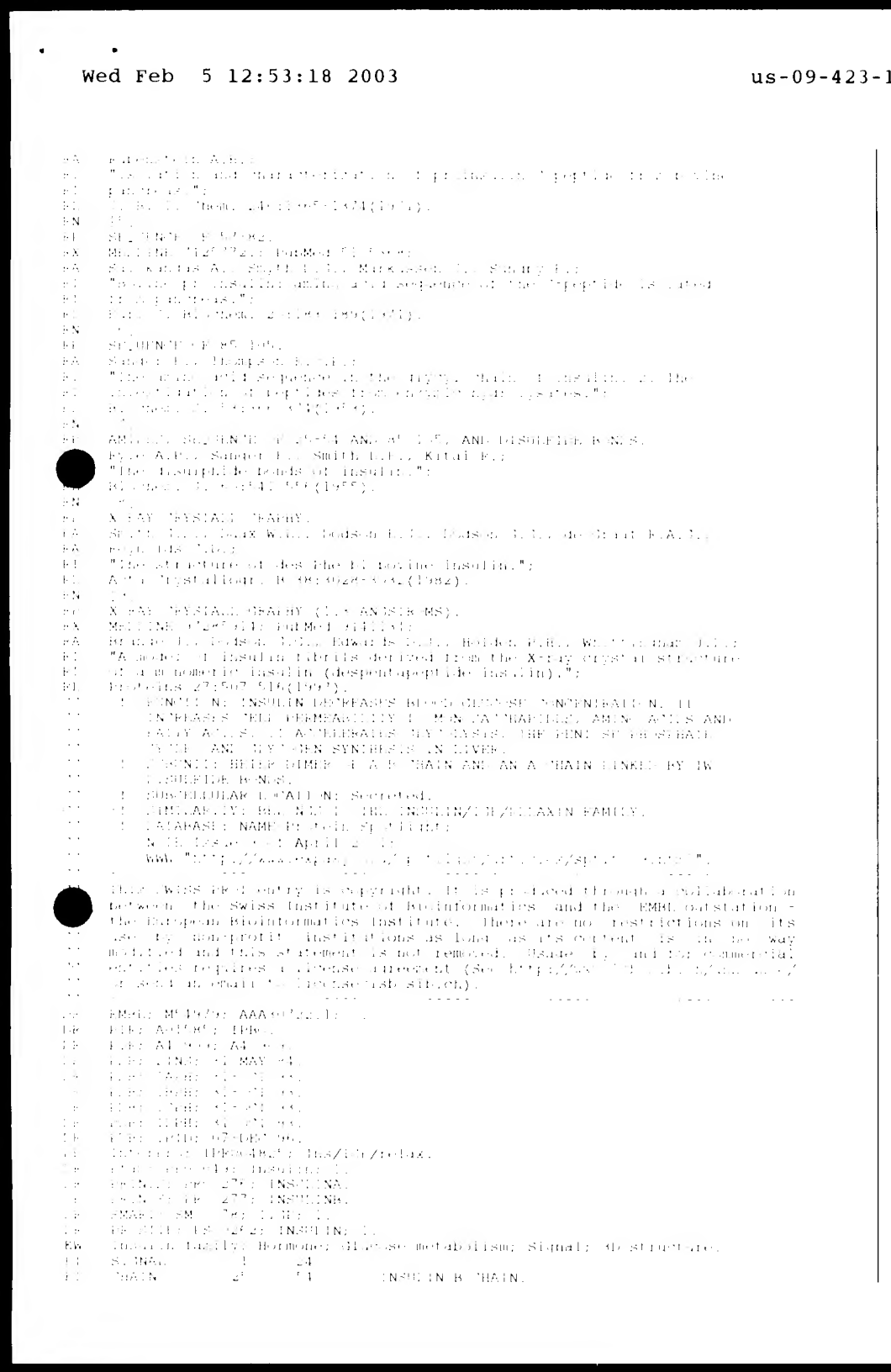
RESULT 1
INS_HUMAN
ID INS_HUMAN STANDARD; PRT; 110 AA.
AC P01308;
DT 21-JUL-1986 (Rel. 1, Created)
DT 21-JUL-1986 (Rel. 2, Last sequence update)
DT 15-JUN-2001 (Rel. 39, Last annotation update)
DE Insulin precursor.
GN INS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=80122[illegible]; PubMed=[illegible];
RA Bell G.I., Pictet R.L., Rutter W.J., Cordell B., Tischer E.,
RA Goodman H.M.;
RT "Sequence of the human insulin gene.";
RL Nature 284:26-32(1980).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=81[illegible]; PubMed=[illegible];
RA Gilbert W., Villa-Komaroff L., Efstratiadis A., Chick W.,
RA Rosenthal N., Jackson R., Kemp D., Lomedico P., Tizard R.,
RA Patarca R., Degen S., Maranto L., Naber S., Kaehler M.,
RA Cohen D., Tizard M.;
RT "Construction and characterization of bacterial plasmids and
RT recombinant plasmids containing initiation regions of rat insulin
RT genes.";
RL Science 209:1171-1176(1980).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=81[illegible]; PubMed=[illegible];
RA Lomedico P., Rosenthal N., Efstratiadis A., Gilbert W.,
RA Kolodner R., Tizard R.;
RT "The structure of the gene for rat insulin.";
RL Cell 18:545-558(1979).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=80[illegible]; PubMed=[illegible];
RA Bell G.I., Swain W.F., Pictet R.L., Cordell B., Goodman H.M.,
RA Rutter W.J.;
RT "Sequence of a human variant insulin gene and detection of a
RT polymorphism in the insulin gene region.";
RL Nature 282:525-527(1979).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=[illegible]; PubMed=[illegible];
RA Soares M.B., Schon E., Henderson A., Karathanasis S.K.,
RA Cate R., Zeitlin S., Chirgwin J., Efstratiadis A.;
RT "RNA-mediated gene duplication: the rat preproinsulin I gene is
RT expressed and has no introns.";
RL Mol. Cell. Biol. 5:2090-2103(1985).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=[illegible]; PubMed=[illegible];
RA Lomedico P., Chirgwin J., Wallace R., Miller J., Gilbert W.,
RA Rutter W.J., Tizard R.;
RT "The structure of the gene for rat proinsulin II.";
RL Cell 13:465-476(1978).
RN [7]
RP SEQUENCE FROM N.A.
RX MEDLINE=[illegible]; PubMed=[illegible];
RA Soares M.B., Schon E., Henderson A., Karathanasis S.K.,
RA Cate R., Zeitlin S., Chirgwin J., Efstratiadis A.;
RT "RNA-mediated gene duplication: the rat preproinsulin I gene is
RT expressed and has no introns.";
RL Mol. Cell. Biol. 5:2090-2103(1985).
RN [8]
RP SEQUENCE FROM N.A.
RX MEDLINE=[illegible]; PubMed=[illegible];
RA Efstratiadis A., Kafatos F.C., Maniatis T.;
RT "The primary structure of rabbit beta-globin gene.";
RL Cell 7:279-288(1976).

RT "[Structure of the 52 amino acid of human insulin]."
RL Submitted (DEC 1998) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE OF 25-54 AND 90-110.
RA Nishi T., Shii K., Hirai J.;
RT "Amino acid sequence of human insulin."
RL Nat. Sci. Rep. Ochanomizu Univ. 19:39-46(1968).
RN [5]
RP SEQUENCE OF 57-87.
RX MEDLINE; 71264404; PubMed 5566404;
RA Ramachandran L.K., Katsoyannis P.G.;
RT "The complete amino acid sequence of the A-peptide of human proinsulin.";
RL Eur. J. Biochem. 20:190-194(1971).
RN [6]
RP SYNTHESIS.
RX MEDLINE; 74077277; PubMed 4443295;
RA Zahn H., Kamber B., Hartmann A., Jost, A., Riniker B., Kittel W.;
RT "Total synthesis of human insulin. Under directed formation of the
RT disulfide bonds.";
RL Helv. Chim. Acta 57:2617-2621(1974).
RN [7]
RP SYNTHESIS OF 57-87.
RX MEDLINE; 75047107; PubMed 4063314;
RA Nicoll D.W.;
RT "Studies on polypeptides. IV. The synthesis of C peptide of human
RT proinsulin.";
RL Hoppe Seyler's Z. Physiol. Chem. 354:669-672(1973).
RN [8]
RP SYNTHESIS OF 25-54 AND 70-87.
RX MEDLINE; 74161261; PubMed 4604553;
RA Schaer H., Hoener J., Kroeni W., Trouet J.;
RT "Synthesis of peptides with the properties of human proinsulin C
RT peptides (hC peptide). 4. Synthesis of the sequences 14-17 and 8-14
RT of human proinsulin C peptides.";
RL Helv. Chim. Acta 56:1-5(1973).
RN [9]
RP SYNTHESIS OF 84-87.
RX MEDLINE; 74161261; PubMed 4604553;
RA Schaer H., Hoener J., Kroeni W., Trouet J.;
RT "Synthesis of peptides with the properties of human proinsulin C
RT peptides (hC peptide). 1. Scheme for the synthesis and preparation of
RT the sequence 28-31 of human proinsulin C peptide.";
RL Chem. Ber. 106:188-192(1973).
RN [10]
RP VARIANT LOS ANGELES.
RX MEDLINE; 74018455; PubMed 4611455;
RA Haneda M., Chan S.J., Kwok S.C.M., Rubenstein A.H., Steiner D.F.;
RT "Studies on mutant human insulin genes: identification and sequence
RT analysis of a gene encoding [SerB24]insulin.";
RL Proc. Natl. Acad. Sci. U.S.A. 72:4824-4828(1975).
RN [11]
RP VARIANTS LOS ANGELES AND CHICAGO.
RX MEDLINE; 74171555; PubMed 4543713;
RA Shoelson S., Haneda M., Blix P., Nanjo K., Sanke T., Inouye K.,
RA Steiner D.F., Rubenstein A.H., Tager H.;
RT "Three mutant insulins in man.";
RL Nature 254:713-716(1975).
RN [12]
RP VARIANT PROVIDENCE.
RX MEDLINE; 87175948; PubMed 3470783;
RA Chan S.J., Seino S., Gruppuso P.A., Schwartz E., Steiner D.F.;
RT "A mutation in the B-chain leader region is associated with impaired
RT pro-insulin conversion in a family with hyperproinsulinemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:7194-7198(1987).
RN [13]
RP VARIANT WAKAYAMA.
RX MEDLINE; 87[illegible]; PubMed [illegible];
RA Nakata H., Iwamoto Y., Sakamoto Y., Kosaka K., Hirata H.,
RT "Structurally abnormal insulin in a diabetic patient. Characterization
RT of the mutant insulin A8 (Val----Leu) isolated from the pancreas.";
RL J. Clin. Invest. 79:[illegible](1987).
RN [14]
RP VARIANT ASP-B3.
RX MEDLINE; 89[illegible]; PubMed [illegible];
RA Barbetti F., Raben N., Kadowaki T., Ohmi A., Accili D., Gabry K.E.,
RA Meredith V.A., Taylor S.I., Roth J.;
RT "Two unrelated patients with functional hyperproinsulinemia due to a
RT naturally occurring substitution of aspartic acid for glycine at
RT position -3 in the insulin presequence identified by direct
RT sequencing of genomic deoxyribonucleic acid amplified by polymerase
RT chain reaction.";
RL J. Clin. Endocrinol. Metab. 71:[illegible](1990).
RN [15]
RP VARIANT ALA-B25.
RX MEDLINE; 90[illegible]; PubMed [illegible];
RA Shibasaki Y., Kawakami T., Kanazawa Y., Akanuma Y., Takaku F.,
RT "Post-translational cleavage of proinsulin is blocked by a point
RT mutation in familial hyperproinsulinemia.";
RL J. Clin. Invest. [illegible](1990).
RN [16]
RP VARIANT KYOTO.
RX MEDLINE; 92291497; PubMed 1601997;
RA Yano H., Kitano N., Morimoto M., Nishinaka K.S., Imura H., Seino Y.,
RT "A novel point mutation in the human insulin gene in a patient with
RT hyperproinsulinemia (proinsulin Kyoto).";
RL J. Clin. Invest. 89:1967-1972(1992).
RN [17]
RP STRUCTURE BY NMR.
RX MEDLINE; 91194966; PubMed 2271664;
RA Hua Q.-X., Weiss M.A.;
RT "Toward the solution structure of human insulin sequential
RT assignment of a des pentapeptide analogue and comparison with crystal
RT structure.";
RL Biochemistry 29:1654-1664(1990).
RN [18]
RP STRUCTURE BY NMR.
RX MEDLINE; 91243747; PubMed 2019753;
RA Hua Q.-X., Weiss M.A.;
RT "Comparative 2D NMR studies of human insulin and des pentapeptide
RT insulin: sequential resonance assignment and implications for protein
RT dynamics and receptor interaction.";
RL Biochemistry 30:5505-5515(1991).
RN [19]
RP STRUCTURE BY NMR.
RX MEDLINE; 91265027; PubMed 1646045;
RA Hua Q.-X., Weiss M.A.;
RT "Two-dimensional NMR studies of Ins (B2 Pr Glu)-Ins: sequence-
RT specific resonance assignments and effects of secondary sequence on
RT [illegible]";
RL Biochim. Biophys. Acta 1079:134-142(1991).
RN [20]
RP STRUCTURE BY NMR.
RX MEDLINE; 95076005; PubMed 7493018;
RA Derewenda U., Derewenda Z., Dodson E.J., Dodson G.G., Reinhart D.,
RT "X-ray analysis of the single-chain B27-insulin mini-B(Asp) mutant at
RT 1.9 A resolution reveals a novel interaction at the carboxy terminus
RT of the insulin B chain.";
RL J. Mol. Biol. 241:760-772(1994).
RN [21]
RP STRUCTURE BY NMR OF VARIANT ILE-A8 ANALOG.
RX MEDLINE; 95[illegible]; PubMed 7644535;
RA Hua Q.-X., Shoelson S.E., Inouye K., Weiss M.A.,
RT "Paradoxical structure and function in a mutant human insulin
RT associated with diabetes mellitus.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:1464-1468(1995).
RN [22]
RP STRUCTURE BY NMR.
RX MEDLINE; 96[illegible]; PubMed 8662707;
RA Chang X., Jorgensen A.M., Bardrum P., Led J.J.,

```
ID   INS_MACFA      STANDARD;      PRT;   110 AA.
AC   P01320; Q04830;
DT   01-JUL-1989 (Rel. 07, Created)
DT   01-JUL-1989 (Rel. 07, Last sequence update)
DT   01-FEB-1996 (Rel. 34, Last annotation update)
DE   Insulin precursor.
GN   INS.
OS   Macaca fascicularis (Crab-eating macaque) (Cynomolgus monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9541;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90208174; PubMed=1689423;
RA   Wentworth W., Chamberlain J., Lichtenwalner M., Wentworth P.,
RA   Wentworth E.;
RT   "Use of variable sequence of cDNA coding for preproinsulin from the
RT   primate Macaca fascicularis.";
RL   Biochem. Biophys. Res. Commun. 169:1-6(1990).
CC   -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC       INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC       FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC       CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC   -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC       DISULFIDE BONDS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE INSULIN FAMILY.
CC
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC
DR   EMBL; X56955; AAA31079.1; -.
DR   PIR; A01726; QVHUM.
DR   PDB; 1INS; 01-JUL-89.
DR   HSSP; P01308; 1ATO.
DR   InterPro; IPR000538; Ins/IGF/relax.
DR   Pfam; PF00049; Insulin; 1.
DR   PRINTS; PR00276; INSULINA.
DR   PRINTS; PR00277; INSULINB.
DR   PROSITE; PS00262; INSULIN; 1.
KW   Insulin family; Hormone; Glucose metabolism; Signal.
FT   SIGNAL      1     24
FT   CHAIN       25    54       INSULIN B CHAIN.
FT   PROPEP      55    87       C-PEPTIDE.
FT   CHAIN       88    110      INSULIN A CHAIN.
FT   DISULFID    31    96       INTERCHAIN.
FT   DISULFID    43    108      INTERCHAIN.
FT   DISULFID    91    106
SQ   SEQUENCE   110 AA;  12409 MW;  95A1F5488ECE427C CRC64;

  Query Match             98.2%;  Score 766;  DB 1;  Length 110;
  Best Local Similarity   99.0%;  Pred. No. 5.4e-42;
  Matches   96;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQ 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQ 84

Qy      61 KRGIVEQCCTSICSLYQLENYCN 84
           ||||||||||||||||||||||||
Db      85 KRGIVEQCCTSICSLYQLENYCN 108

ID   INS_MACFA      STANDARD;      PRT;   110 AA.
AC   P01320; Q04830;
DT   01-JUL-1989 (Rel. 07, Created)
DT   01-JUL-1989 (Rel. 07, Last sequence update)
DT   01-FEB-1996 (Rel. 34, Last annotation update)
DE   Insulin precursor.
GN   INS.
OS   Macaca fascicularis (Crab-eating macaque) (Cynomolgus monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9541;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90208174; PubMed=1689423;
RA   Wentworth W., Chamberlain J., Lichtenwalner M., Wentworth P.,
RA   Wentworth E.;
RT   "Use of variable sequence of cDNA coding for preproinsulin from the
RT   primate Macaca fascicularis.";
RL   Biochem. Biophys. Res. Commun. 169:1-6(1990).
CC   -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC       INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC       FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC       CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC   -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC       DISULFIDE BONDS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE INSULIN FAMILY.
CC
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC
DR   EMBL; X56955; AAA31079.1; -.
DR   PIR; A01726; QVHUM.
DR   PDB; 1INS; 01-JUL-89.
DR   HSSP; P01308; 1ATO.
DR   InterPro; IPR000538; Ins/IGF/relax.
DR   Pfam; PF00049; Insulin; 1.
```







Run on: February 5, 2003, 12:42:21 ; Search time 0.352 Seconds
(without alignments)
831.727 Million cell updates/sec

Title: **US-09-423-100-4**
Perfect score: 463

Sequence: 1 GVNQHLCGSHLVEALYLVC.....VEQKCCSICYDGIENYCN 49

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1754 seqs, 1.0 47117 residues

Total number of hits satisfying chosen parameters: 671586

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : **SPTREMBL_21:***
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_phage:*
8: sp_plant:*
9: sp_rodent:*
10: sp_virus:*
11: sp_vertebrate:*
12: sp_unclassified:*
13: sp_viruses:*
14: sp_bacteriap:*
15: sp_archaep:*

No. N: is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match Length	DB	ID	Description
1	114	100.0	110	11	Q91XI3 Spermophi...
2	109	97.9	111	1	Q9WNW7 Ictalurus...
3	107.5	97.9	111	1	Q9I7J7 Danio pir...
4	107.5	97.9	111	1	Q9I7J7 Oreochrom...
5	107.5	97.9	110	1	Q9I7A8 Poecilia...
6	106.5	97.9	110	1	Q9I7Y1 Poecilia...
7	103	97.9	111	1	Q9I7B0 Chitala ...
8	101.5	97.9	109	1	Q9I7E5 Ictalurid...
9	100.5	97.9	1	1	Q9YGN1 Salvelinu...
10	99.5	97.9	1	1	Q9I7A7 Oncorhync...
11	99.5	97.9	1	1	Q9I7B1 Carassius...
12	99.5	91.8	1	1	Q9I7B2 Ambloplit...
13	99	100.0	11	11	Q9Z543 Mus spret...
14	99	97.9	11	11	Q9Z542 Mus spret...
15	98.5	97.9	109	1	Q9I7D7 Paralicht...
16	98.5	97.9	109	1	Q7T9Z8 Oreochrom...
17	98.5	97.9	110	1	Q9I7A4 Poecilia...
18	98.5	97.9	110	1	Q9I7A5 Poecilia...
19	98.5	97.9	110	1	Q9I7A6 Poecilia...
20	98.5	97.9	110	1	Q9I7A9 Poecilia...
21	98	97.9	1	1	Q9YGM8 Salvelinu...
22	97.5	97.9	109	1	Q9I7D6 Paralicht...
23	97.5	97.9	109	1	Q9I7D8 Paralicht...
24	97.5	97.9	110	1	Q9I7A3 Poecilia...
25	96.5	97.9	1	1	Q9YGM9 Salvelinu...
26	96.5	97.9	1	1	Q9YGN0 Salvelinu...
27	96.5	97.9	109	1	Q9I7D9 Paralicht...
28	96.5	97.9	110	1	Q9I7A2 Poecilia...
29	96.5	97.9	110	1	Q9I7B3 Poecilia...
30	96.5	97.9	110	1	Q9I7B4 Poecilia...
31	96	97.9	1	1	Q9YGM7 Salvelinu...
32	95.5	97.9	1	1	Q9YGM6 Salvelinu...
33	95.5	97.9	1	1	Q9YGN2 Salvelinu...
34	95.5	97.9	1	1	Q9YGN3 Salvelinu...
35	95.5	97.9	1	1	Q9YGN4 Salvelinu...
36	95.5	97.9	1	1	Q9YGN5 Salvelinu...
37	95.5	97.9	1	1	Q9YGN6 Salvelinu...
38	95.5	97.9	1	1	Q9YGN7 Salvelinu...
39	95.5	97.9	1	1	Q9YGN8 Salvelinu...
40	95.5	97.9	1	1	Q9YGN9 Salvelinu...
41	95.5	97.9	1	1	Q9YGP0 Salvelinu...
42	95.5	97.9	1	1	Q9YGP1 Salvelinu...
43	95	97.9	1	1	Q9YGP2 Salvelinu...
44	95	97.9	1	1	Q9YGP3 Salvelinu...
45	95	97.9	1	1	Q9YGP4 Salvelinu...

ALIGNMENTS

RESULT 1
Q91XI3
ID Q91XI3 PRELIMINARY; PRT; 110 AA.
AC Q91XI3;
DT 01-DEC-2000 (TrEMBLrel. 14, Created)
DT 01-DEC-2000 (TrEMBLrel. 14, Last sequence update)
DT 15-MAR-2002 (TrEMBLrel. 19, Last annotation update)
DE Ferritin.
OS Spermophilus tridecemlineatus (Thirteen-lined ground squirrel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciuromorpha; Sciuridae; Spermophilus.
OX NCBI_TaxID=43179;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Jenkins M.M., Buck M.J., Schandl C.A., Spurio L.J., Andrews N.C.;
RT "Regulation of HFE expression in a hibernating mammal.";
RL Submitted (JUN-2000) to EMBL/GenBank/DDBJ databases.
CC -!- SUBCELLULAR LOCATION: NOT SPECIFIED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INSECTIN/FERRITIN FAMILY.
DR EMBL; AY049044; AAK71744.1; -.
DR InterPro; IPR002024; Insectin.
DR Pfam; PF00210; Insectin; 1.
DR PROSITE; PS00210; INSECTIN; 1.
SQ SEQUENCE 110 AA; 12355 MW; 5A5A3A5A5A5A5A5A CRC64;

Query Match 97.9%; Score 453; DB 1; Length 110;
Best Local Similarity 97.9%; Pred. No. 1.1e-42;
Matches 48; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCSICYDGIENYCN 80

Db 1 GVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCSICYDGIENYCN 80

Qy 81 SL 82

Db 81 SL 82

RESULT 3
Q9WNW9
ID Q9WNW9 PRELIMINARY; PRT; 11 AA.
AC Q9WNW9;
DT 01-MAR-2000 (TrEMBLrel. 20, Created)
DT 01-MAR-2000 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Proinsulin.
GN Felis silvestris catus (Cat).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PANCREAS;
RA Kato K., Morimatsu M.;
RL "Cat insulin.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AB041545; BAB64110.1; -.
DR InterPro; IPR004825; Ins/Gly/Ovo/Lym;
DR Pfam; PF00049; Insulin; 1.
DR PRINTS; PR00276; INSULINA.
DR PRINTS; PR00277; INSULINB.
DR SMART; SM00078; IlPc; 1.
DR PROSITE; PS00262; INSULIN; UNKNOWN_1.
SQ SEQUENCE 11 AA; 1209 MW; 9FBF8B7B077B87A1 CRC64;

Query Match 90.9%; Score 780; DB 5; Length 11;
Best Local Similarity 94.7%; Pred. No. 1.9e-09;
Matches 72; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKR 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKR 60

Qy 61 GIVEQCCTSICSLYQLENYCN 81
||||||||||||||||||||
Db 61 GIVEQCCTSICSLYQLENYCN 81

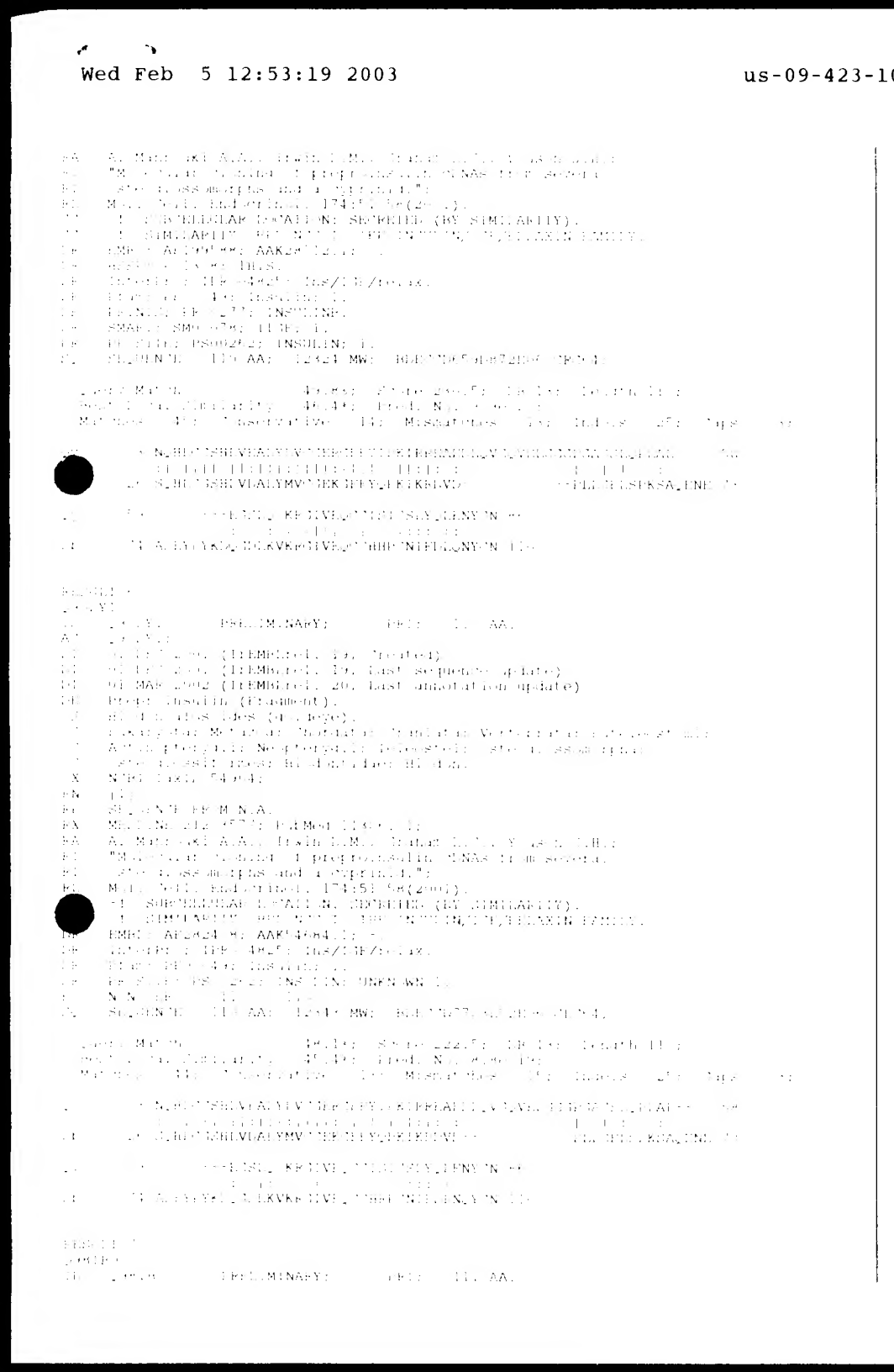
RESULT 4
Q9W1A7
ID Q9W1A7 PRELIMINARY; PRT; 11 AA.
AC Q9W1A7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 21, Last annotation update)
DE Preproinsulin (Fragment).
GN Hydrocharis dubia (Frogbit) (Stratiotes dubia).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
OX NCBI_TaxID=8404;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22125677; PubMed=12019274;
RA Irwin D.M., Sivaraajah P.;
RT "Preproinsulin cDNAs from the leopard frog, Rana pipiens: evolution of
RT proinsulin processing.";
RL Comp. Biochem. Physiol. 129B:615-621(2001).
DR EMBL; AF322147; AAF87285.1; -.
DR InterPro; IPR004825; Ins/Gly/Relaxin FAM;.
DR Pfam; PF00049; Insulin; 1.
DR PRINTS; PR00276; INSULINA.
DR PRINTS; PR00277; INSULINB.
DR SMART; SM00078; IlPc; 1.
DR PROSITE; PS00262; INSULIN; 1.
SQ SEQUENCE 11 AA; 1210 MW; 2A97DE8977A217F02 CRC64;

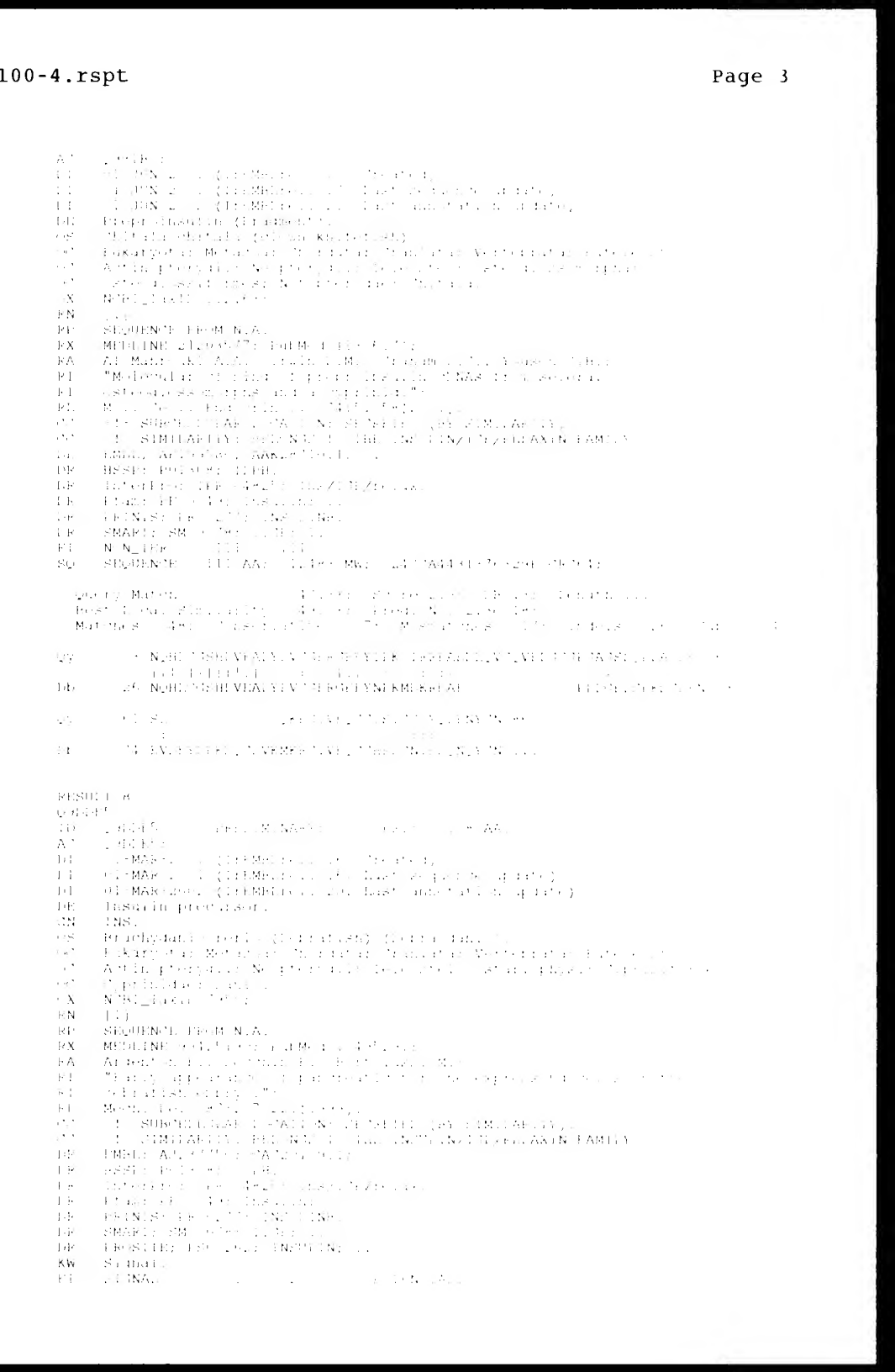
Query Match 90.9%; Score 780; DB 5; Length 11;
Best Local Similarity 94.7%; Pred. No. 1.9e-09;
Matches 72; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKR 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 FVNQHLCGSHLVEALYMVCGDRGFFYSPKSRREVEDPQVEQLELGGSPGDLQTLALEVAQQKR 60

Qy 61 GIVEQCCTSICSLYQLENYCN 81
||||||||||||||||||||
Db 61 GIVEQCCASVCSLYQLENYCN 81

RESULT 5
Q9W1A8
ID Q9W1A8 PRELIMINARY; PRT; 110 AA
AC Q9W1A8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 21, Last annotation update)
DE Preproinsulin.
GN Pantodon buchholzi (Butterfish).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossidae; Pantodontidae; Pantodon.
OX NCBI_TaxID=8376;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22125677; PubMed=12019274;





```
RESULT 12
Q08182
ID   Q08182       PRELIMINARY;      PRT:    41 AA.
AC   Q08182;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Preproinsulin (Fragment).
OS   Amblyopsis spelaea.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidei;
OC   Centrarchidae; Amblopliti.
OX   NCBI_TaxID=6973;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Al-Mahrouki A.A., Irwin L.M., Youson J.H.;
RL   "Molecular cloning of preproinsulin cDNA from the cave bass.";
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE INSULIN FAMILY.
DR   EMBL; AF159081; AAK26708.1; -.
DR   HSSP; P01308; 1TPH.
DR   InterPro; IPR004061; Insulin_C.
DR   Pfam; PF00049; Insulin; 1.
DR   PRINTS; PR00276; INSULINA.
DR   SMART; SM00078; IlGF; 1.
DR   PROSITE; PS00262; INSULIN; 1.
KW   Hormone; Glucose metabolism; Multigene family.
FT   NON_TER       1       1
FT   CHAIN         21      41      A CHAIN.
SQ   SEQUENCE   41 AA;  4641 MW;  [illegible] CRC64;

  Query Match             44.0%;  Score 20.5;  DB 13;  Length 41;
  Best Local Similarity   46.7%;  Pred. No. 1.5e-16;
  Matches   42;  Conservative   13;  Mismatches   26;  Indels    9;  Gaps    4;

Qy       3 [illegible] 59
           |||||| ||||||||||||| || ||  |  |  ||  |  |  |
Db       4 [illegible] 42

Qy      60 [illegible] 86
           ||  ||||||||  |||  |  |||
Db      43 [illegible] 69


RESULT 13
Q62542
ID   Q62542       PRELIMINARY;      PRT:    41 AA.
AC   Q62542;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Insulin 2 (Fragment).
GN   INS2.
OS   Mus spretus (Western wild mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10096;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   STRAIN=SPRET/EI;
RX   MEDLINE=94319082; PubMed=8041648;
RA   Bonner-Weir S., Weir G.C., Burke G.T., Baxter M.L., Takahashi N., Nazzaki Y.,
RA   Nadeau J.H.;
RT   "Genetic mapping of 40 cDNA clones on the mouse genome by PCR.";
RL   Mamm. Genome 5:349-355(1994).
CC   -!- FUNCTION: INSULIN HAS PLEIOTROPIC REGULATORY FUNCTIONS. IT
CC       INCREASES THE PERMEABILITY OF MANY CELLS TO MANY METABOLITES, AMONG
CC       FATTY ACIDS, IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC       CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC   -!- SUBUNIT: HEXAMER: 2 A CHAIN AND 2 B CHAIN TWO
CC       DISULFIDE BONDS.
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR   EMBL; U05729; AAA50473.1; -.
DR   HSSP; P01322; 1A7F.
DR   MGD; MGI:96572; Ins2.
DR   InterPro; IPR004061; Ins/IGF/relax.
DR   Pfam; PF00049; Insulin; 1.
DR   PRINTS; PR00276; INSULINA.
DR   SMART; SM00078; IlGF; 1.
DR   PROSITE; PS00262; INSULIN; 1.
KW   Hormone; Glucose metabolism; Multigene family.
FT   NON_TER       1       1
FT   CHAIN         21      41      A CHAIN.
SQ   SEQUENCE   41 AA;  4641 MW;  [illegible] CRC64;

  Query Match             44.0%;  Score 20.5;  DB 13;  Length 41;
  Best Local Similarity   46.7%;  Pred. No. 1.5e-16;
  Matches   42;  Conservative   13;  Mismatches   26;  Indels    9;  Gaps    4;

Qy       3 [illegible] 59
           |||||| ||||||||||||| || ||  |  |  ||  |  |  |
Db       4 [illegible] 42

Qy      60 [illegible] 86
           ||  ||||||||  |||  |  |||
Db      43 [illegible] 69


RESULT 14
Q62542
ID   Q62542       PRELIMINARY;      PRT:    41 AA.
AC   Q62542;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Insulin 2 (Fragment).
GN   INS2.
OS   Mus spretus (Western wild mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10096;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   STRAIN=SPRET/EI;
RX   MEDLINE=94319082; PubMed=8041648;
RA   Bonner-Weir S., Weir G.C., Burke G.T., Baxter M.L., Takahashi N., Nazzaki Y.,
RA   Nadeau J.H.;
RT   "Genetic mapping of 40 cDNA clones on the mouse genome by PCR.";
RL   Mamm. Genome 5:349-355(1994).
CC   -!- FUNCTION: INSULIN HAS PLEIOTROPIC REGULATORY FUNCTIONS. IT
CC       INCREASES THE PERMEABILITY OF MANY CELLS TO MANY METABOLITES, AMONG
CC       FATTY ACIDS, IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC       CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC   -!- SUBUNIT: HEXAMER: 2 A CHAIN AND 2 B CHAIN TWO
CC       DISULFIDE BONDS.
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR   EMBL; U05729; AAA50473.1; -.
DR   HSSP; P01322; 1A7F.
DR   MGD; MGI:96572; Ins2.
DR   InterPro; IPR004061; Ins/IGF/relax.
DR   Pfam; PF00049; Insulin; 1.
DR   PRINTS; PR00276; INSULINA.
DR   SMART; SM00078; IlGF; 1.
DR   PROSITE; PS00262; INSULIN; 1.
KW   Hormone; Glucose metabolism; Multigene family.
FT   NON_TER       1       1
FT   CHAIN         21      41      A CHAIN.
SQ   SEQUENCE   41 AA;  4641 MW;  [illegible] CRC64;

  Query Match             44.0%;  Score 20.5;  DB 13;  Length 41;
  Best Local Similarity   46.7%;  Pred. No. 1.5e-16;
  Matches   42;  Conservative   13;  Mismatches   26;  Indels    9;  Gaps    4;

Qy       3 [illegible] 59
           |||||| ||||||||||||| || ||  |  |  ||  |  |  |
Db       4 [illegible] 42

Qy      60 [illegible] 86
           ||  ||||||||  |||  |  |||
Db      43 [illegible] 69
```

```
DB       1 GSPGNDAPTALDVAALQKEGIVEECCFRSCDLALLETYCATPAKSE 49


RESULT 15
O93607
ID   O93607      PRELIMINARY;      PRT:   159 AA.
AC   O93607;
DT   01-NOV-1998 (TrEMBLrel. 03, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Preproinsulin-like growth factor Ia.
GN   IGF-I.
OS   Paralichthys olivaceus (Flounder).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC   Pleuronectoidei; Paralichthyidae; Paralichthys.
OX   NCBI_TaxID=8255;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kim I.S.;
RT   "Expression of IGF Ib cDNA clone isolated from Paralichthys olivaceus
RT   in mammalian CHO cell line using green fluorescence protein (GFP)
RT   tagging: secretory production of big IGFIb-GFP fusion proteins from
RT   stable transfected CHO cell culture.";
RL   Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE [illegible] FAMILY.
DR   EMBL; AJ010602; CAA09267.1; -.
DR   BSSP; P01343; IGF1.
DR   InterPro; IPR004825; Ins/IGF/relax.
DR   Pfam; PF00049; Insulin; 1.
DR   PRINTS; PR00276; INSULINA.
DR   PRINTS; PR00277; INSULINB.
DR   SMART; SM00078; IlGF; 1.
DR   PROSITE; PS00262; INSULIN; 1.
SQ   SEQUENCE   159 AA;  17511 MW;  3B9C1D7D9D1E2C85 CRC64;

  Query Match             30.8%;  Score 142.5;  DB 13;  Length 159;
  Best Local Similarity   37.5%;  Pred. No. 4.3e-09;
  Matches   30;  Conservative    9;  Mismatches   16;  Indels   2%;  Gaps    2;

Qy       6 DPGSHLVLALYLACQGGFFYTPKTGYGSSSRRAPQTGIVDECCFRSCDLALLETYCATPAKSE 65
           ||| || | |||||||:::  |               | || |             : |
Db      49 DPGSEAVFTDLDVVSQDFKQYLQQVRAENEDGSLQSDQFSFQFFNLNSIKSPEERDFISQVV 108

Qy      66 GIVEQCCTSICSLYQLENYC 85
           :|:::|:   | | :|| ||
Db     109 GIVDECCFQSCELRRLEMYC 128


Search completed: February  5, 2003, 12:47:29
Job time : 28.3502 secs
```

Title: US-09-423-100-4
Perfect score: 463

Database : Issued_Patents_AA:*

SUMMARIES

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Software #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/423,100
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Leslie McDonell
; REGISTRATION NUMBER: 34,082
; REFERENCE/DOCKET NUMBER: 0281.1000-000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: Site
; LOCATION: 1..9
; OTHER INFORMATION: /note= "Xaa = Ala or Val"

Query Match 100.0%; Score 40; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.3e-47;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EVNLRSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 EVNLRSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60

Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 86
|||||||||||||||||||||||||||
Db 61 SLQKRGIVEQCCTSICSLYQLENYCN 86

RESULT 1
US-08-160-276A-4
; Sequence 4, Application US/08160276A
; Patent No. 5476936
; GENERAL INFORMATION:
; APPLICANT: Goldsmith, Randy
; APPLICANT: Gerl, Martin
; APPLICANT: Ludwig, Jurgen
; APPLICANT: Sabel, Walter
; TITLE OF INVENTION: Expression of Foreign Genes in
; TITLE OF INVENTION: Spodoptera Cells
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive, 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH, 1.44 MB STORAGE
; COMPUTER: IBM PC
; OPERATING SYSTEM: WINDOWS 3.1
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/160,276A
; FILING DATE: December 1, 1993
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/919,914
; FILING DATE: 24-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Barbara V. Mantei, Esq.
; REGISTRATION NUMBER: 31,620
; REFERENCE/DOCKET NUMBER: B-E 276 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (714) 760-1020
; TELEFAX: (714) 760-2187

```
;    APPLICANT:  Nowicki-Washington, Karl D.
;    APPLICANT:  Clark, Samuel A.
;    APPLICANT:  Thigpen, Alice E.
;    APPLICANT:  Quaade, Christian
;    APPLICANT:  Krause, Fred
;    APPLICANT:  McGarry, Dennis
;    TITLE OF INVENTION:  RECOMBINANT EXPRESSION OF PROTEINS FROM
;    TITLE OF INVENTION:  SECRETORY CELL LINES
;    NUMBER OF SEQUENCES:  79
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  [illegible]
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/784,582
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US [illegible]/[illegible],427
;      FILING DATE:  [illegible]
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/585,028
;      FILING DATE:  19-JAN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Highlander, Steven L.
;      REGISTRATION NUMBER:  37,642
;      REFERENCE/DOCKET NUMBER:  UTSD:514
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/418-3000
;      TELEFAX:  512/474-7577
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  110 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-784-582-2

  Query Match             [illegible];  Score 4[illegible];  DB [illegible];  Length 110;
  Best Local Similarity   [illegible];  Pred. No. 1.5e-47;
  Matches   86;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
Db         25 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 84

Qy         61 SLQKRGIVEQCCTSICSLYQLENYCN 86
Db         85 SLQKRGIVEQCCTSICSLYQLENYCN 110


RESULT 8
US-08-784-582-2
; Sequence 2, Application US/08784582
; Patent No. 6004170
;  GENERAL INFORMATION:
;    APPLICANT:  Newgard, Christopher B.
;    APPLICANT:  Halban, Phillippe A.
;    APPLICANT:  Nowicki-Washington, Karl D.
;    APPLICANT:  Clark, Samuel A.
;    APPLICANT:  Thigpen, Alice E.
;    APPLICANT:  Quaade, Christian
;    APPLICANT:  Krause, Fred
;    TITLE OF INVENTION:  RECOMBINANT EXPRESSION OF PROTEINS FROM
;    TITLE OF INVENTION:  SECRETORY CELL LINES
;    NUMBER OF SEQUENCES:  79
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/785,271
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/585,028
;      FILING DATE:  19-JAN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Highlander, Steven L.
;      REGISTRATION NUMBER:  37,642
;      REFERENCE/DOCKET NUMBER:  UTSD:514
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/418-3000
;      TELEFAX:  512/474-7577
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  110 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-785-271-2

  Query Match             [illegible];  Score 4[illegible];  DB 4;  Length 110;
  Best Local Similarity   [illegible];  Pred. No. 1.5e-47;
  Matches   86;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
Db         25 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 84

Qy         61 SLQKRGIVEQCCTSICSLYQLENYCN 86
Db         85 SLQKRGIVEQCCTSICSLYQLENYCN 110


RESULT 9
PCT-US95-08596-2
; Sequence 2, Application PC/TUS9508596
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Proinsulin Peptide Compounds for Detecting
;    TITLE OF INVENTION:  and Treating Type I Diabetes
;    NUMBER OF SEQUENCES:  [illegible]
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 State Street, suite 510
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109-1875
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/08596
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
```

; REFERENCE/APPLICATION NUMBER: [illegible]
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408 4400
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..99
US-09-423-100-4

Query Match 98.7%; Score 497; DB 2; Length 99;
Best Local Similarity 100.0%; Pred. No. 4.6e-47;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
Db 15 PVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 74

Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 85
Db 75 SLQKRGIVEQCCTSICSLYQLENYCN 99

RESULT 5
US-09-423-100-5
; Sequence 5, Application US/09423100A
; Patent No. 6380365
; GENERAL INFORMATION:
; APPLICANT: Rubinoder, Franz-Josef
; Keller, Reinhold
; TITLE OF INVENTION: Improved process for obtaining
; insulin precursors having correctly bonded cystine bridges
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/423,100A
; FILING DATE: 01-Apr-1999
; CLASSIFICATION: Unknown
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,856
; FILING DATE: Unknown
; ATTORNEY/AGENT INFORMATION:
; NAME: Leslie McDonald
; REGISTRATION NUMBER: 34,472
; REFERENCE/DOCKET NUMBER: 0248.1600-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408 4400
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..99
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-423-100-5

Query Match [illegible]; Score [illegible]; DB 4; Length 99;
Best Local Similarity [illegible]; Pred. No. [illegible];
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy 1 PVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
[illegible]
Db 15 PVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 74

Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 85
[illegible]
Db 75 SLQKRGIVEQCCTSICSLYQLENYCN 99

Search completed: February [illegible]
Job time : 15.6688 secs

Title: US-09-423-100-4
Perfect score: 463

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*

SUMMARIES

ALIGNMENTS

; APPLICANT: Dunaiski, Yvonne
; APPLICANT: Joubert, Dirk
; TITLE OF INVENTION: DP-IBIN Variants
; FILE REFERENCE: 117181.1
; CURRENT APPLICATION NUMBER: US/10/028,411
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US/09/477,924
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ IDS: 9
; SEQ ID NO 1
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-028-411-2

Query Match 100.0%; Score 46; DB 4; Length 86;
Best Local Similarity 100.0%; Pred. No. 4.3e-44;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EVNLRLCVSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
Db 1 EVNLRLCVSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60

Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 86
Db 61 SLQKRGIVEQCCTSICSLYQLENYCN 86

RESULT 4
US-10-028-411-4
; Sequence 4, Application US/10028411
; Patent No. US20020147412A1
; GENERAL INFORMATION:
; APPLICANT: Tan, Thong Pi
; TITLE OF INVENTION: Chimeric Protein Containing an
; Intramolecular Chaperone Like Sequence
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/028,411
; FILING DATE: 21-Dec-2001
; CLASSIFICATION: Unknown
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/477,924
; FILING DATE: 05-JAN-2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers II, Frank C
; REGISTRATION NUMBER: 46,949
; REFERENCE/DOCKET NUMBER: 020167-001300US
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 86 amino acids
; TYPE: amino acid
; STRANDEDNESS: Unknown
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-028-411-4

Query Match 100.0%; Score 46; DB 4; Length 86;
Best Local Similarity 100.0%; Pred. No. 4.3e-44;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EVNLRLCVSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
Db 1 EVNLRLCVSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60

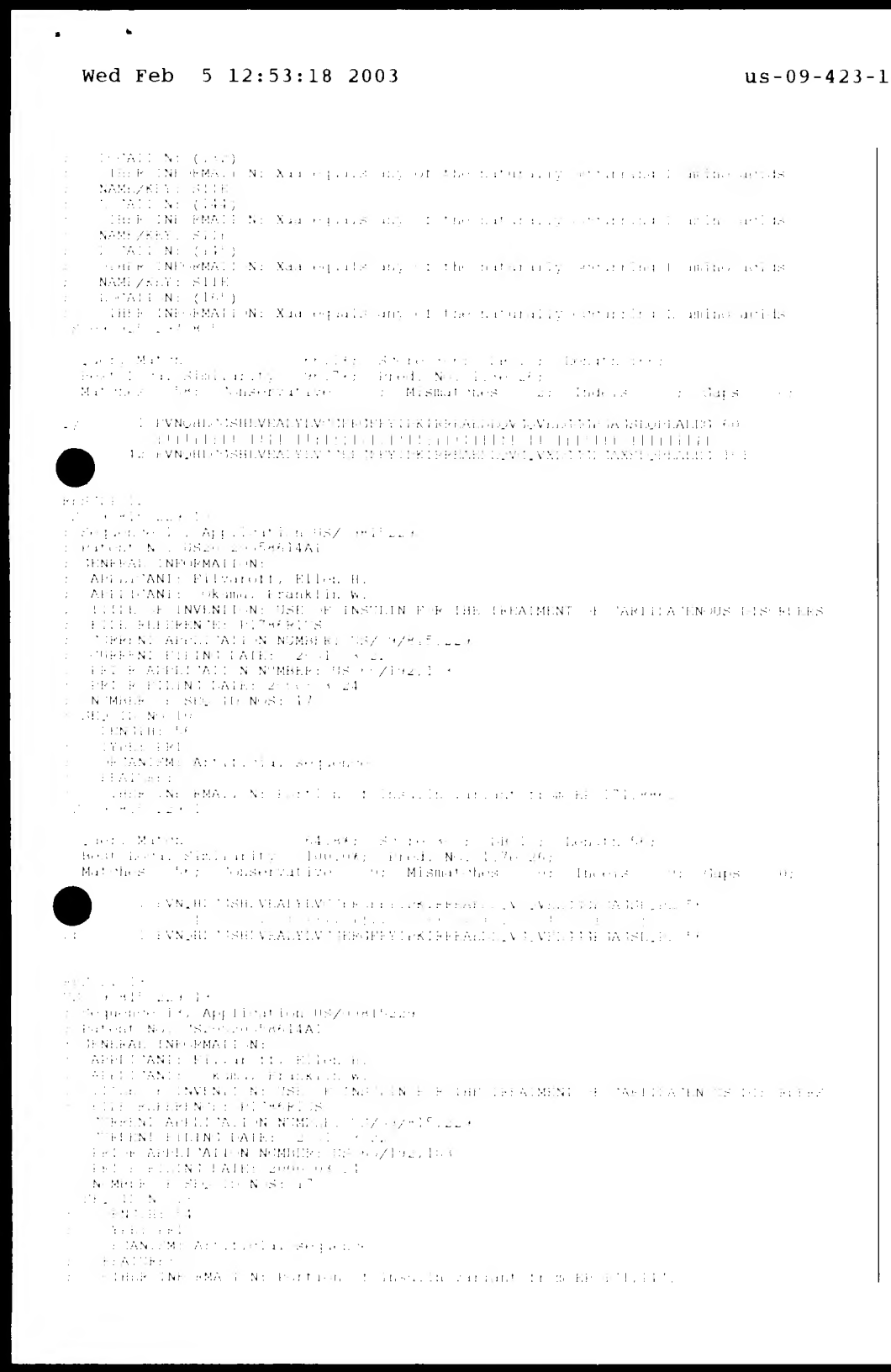
Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 86
Db 61 SLQKRGIVEQCCTSICSLYQLENYCN 86

RESULT 4
US-10-028-411-4
; Sequence 4, Application US/10028411
; Patent No. US20020147412A1
; GENERAL INFORMATION:
; APPLICANT: Rubrecht, Franc Josef
; Keller, Reinhold
; TITLE OF INVENTION: Type 2 Diabetes Related Gene
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/944,583
; FILING DATE: 07-Sep-2001
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/477,924
; FILING DATE: 05-Jan-2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Leslie Mehaffie
; REGISTRATION NUMBER: 41,605
; REFERENCE/DOCKET NUMBER: 10759-[illegible]
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..9
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-944-583-4

Query Match 100.0%; Score 46; DB 4; Length 86;
Best Local Similarity 100.0%; Pred. No. 4.3e-44;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EVNLRLCVSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
Db 1 EVNLRLCVSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60

Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 86



Title: US-09-423-100-5
Perfect score: 294

Database : A_Geneseq_101002:*

SUMMARIES

Description

Human insulin prec
Human pre-insulin
Bovine pre-insulin
Human pre-insulin
Fusion protein con
hcIB mini-proinsuli
Sfd proinsulin hyb
Matrix factor alph
Matrix factor alph

ALIGNMENTS

```
PS   Claim 12; Page 29-30; 46pp; English.
XX
CC   This sequence represents a human insulin precursor comprising insulin
CC   A and B chains. This insulin precursor is a component of the chimeric
CC   proteins hGH-mini-proinsulin (AAY42866) and the chimeric protein
CC   given in AAY42951. These chimeric proteins additionally contain an
CC   N-terminal fragment of human growth hormone (hGH) and a cleavable
CC   peptide linker (AAY12957). The hGH portion of the chimeric protein acts
CC   as an intramolecular chaperone (IMC) for the insulin precursor,
CC   enabling it to fold correctly. The cleavable peptide linker has a
CC   C-terminal Arg residue which enables the hGH portion of the
CC   chimeric protein to be removed after folding has taken place. Production
CC   of recombinant human insulin via an hGH-proinsulin chimeric protein can
CC   provide human insulin with correctly linked cystine bridges with
CC   fewer necessary procedural steps, and hence resulting in a higher yield
CC   of human insulin. The IMC sequences not only protect insulin sequences
CC   from intracellular degradation by a microorganism host, but also promote
CC   the folding of the fused insulin precursor, facilitate the solubility of
CC   the fusion protein and decrease the intermolecular interactions among
CC   the fusion proteins, thus allowing folding of the fused insulin precursor
CC   at commercially useful high concentrations. The procedural steps of
CC   cyanogen bromide cleavage, oxidative sulphitolysis and related
CC   purification steps can thus be eliminated, along with the use of high
CC   concentrations of mercaptan or the use of hydrophobic absorbent resins.
XX
SQ   Sequence   52 AA;

  Query Match             100.0%;  Score 294;  DB 20;  Length 52;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-26;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREVEDPQVEQLELGGSPGDLQTLALEVAQQKRGIVDQCCTSICSLYQLENYCN 52
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREVEDPQVEQLELGGSPGDLQTLALEVAQQKRGIVDQCCTSICSLYQLENYCN 52


RESULT 2
AAR98666
ID   AAR98666 standard; peptide; 56 AA.
XX
AC   AAR98666;
XX
DT   02-MAR-1995  (first entry)
XX
DE   Human pro-insulin 3.
XX
KW   Pro-insulin; A chain; B chain; C chain; disulphide;
KW   mercaptan; chaotropic agent.
XX
OS   Homo sapiens.
XX
PN   EP608292-A2.
XX
PD   03-AUG-1994.
XX
PF   25-NOV-1993;  93EP-0118993.
XX
PR   02-DEC-1992;  92DE-4240420.
XX
PA   (BASF ) BASF SYST AG.
XX
PI   Gerl M,  Lindwin J,  Obermeier R,  Sabel W;
XX
DR   WPI; 1994-177716/22
XX
PT   Prepn. of pro-insulin with correct di-sulphide bridges - by
PT   treating recombinant precursor protein with mercaptan in alkali
PT   and in presence of chaotropic agent, then isolation on
PT   hydrophobic resin
XX
PS   Disclosure; Page 12; 15pp; German.
XX
CC   Pro-insulin is produced by treating recombinant precursor protein
CC   with a mercaptan to provide 2 to 50 SH residues per Cys residue, in
CC   presence of a chaotropic agent and in aq. medium of pH 10-11,
CC   treating the prod. with 3-50 x hydrophobic adsorber resin per l. aq.
CC   medium of pH 4-7, isolating the adsorbed resin and pro-insulin, and
CC   desorbing the pro-insulin. This method produces pro-insulin with
CC   correctly banded Cys bridges. Compared with known methods it
CC   involves fewer stages (esp. no sulphitolysis or cyanogen bromide
CC   cleavage) and overall losses during purification are reduced, i.e.
CC   the process is quicker and gives better yields.
CC   Sequences of insulin chain A, B and C are given in AAR68897-99.
CC   Sequences of pro-insulin 1-4 are given in AAR68898-901.
XX
SQ   Sequence   56 AA;

  Query Match             100.0%;  Score 294;  DB 15;  Length 56;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-26;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREVEDPQVEQLELGGSPGDLQTLALEVAQQKRGIVDQCCTSICSLYQLENYCN 52
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        5 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREVEDPQVEQLELGGSPGDLQTLALEVAQQKRGIVDQCCTSICSLYQLENYCN 56


RESULT 3
AAR78665
ID   AAR78665 standard; protein; 56 AA.
XX
AC   AAR78665;
XX
DT   03-APR-1996  (first entry)
XX
DE   Proinsulin sequence 3.
XX
KW   Proinsulin; post-translational modification; recombinant protein; c.;
KW   protein folding; conformation.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Region          1..4
FT                   /label= B2
FT                   /note= "a peptide of 4 amino acids"
FT   Peptide         5..34
FT                   /label= B1-(B2-B29)-Y
FT                   /note= "human insulin B-chain"
FT   Region          35
FT                   /label= X
FT   Peptide         36..56
FT                   /label= CHY-(A2-A20)-E3
FT                   /note= "human insulin A-chain"
XX
PN   EP608292-A2.
XX
PD   23-AUG-1994.
XX
PF   09-FEB-1994;  94EP-0101748.
XX
PR   19-FEB-1993;  93DE-4305179.
XX
PA   (BASF ) BASF AG.
XX
PI   Gerl M,  Lindwin J,  Obermeier R,  Sabel W;
XX
DR   WPI; 1994-263756/32.
XX
PT   Isolation of insulin that is correctly post-translationally
PT   processed - by reacting proinsulin with a mercaptan in the presence
PT   of a chaotropic agent and partition, after absorption to hydrophobic
PT   resin
XX
PS   Example 2; Page 14; 16pp; German.
XX
CC   The present sequence is an example of a proinsulin molecule correct.
```

RESULT 1
AAB70012
LC AAB70012 standard; protein; 86 AA.
XX
AC AAB70012;
XX
DT 15-JAN-2000 (first entry)
XX
DE Basic protein conju. proinsulin sequence 4.
XX
KW Proinsulin; post-translational modification; recombinant production;
KW protein folding; conformation.
XX
OS Synthetic.
XX
FH Key Location/qualifiers
FT Peptide 1..44
FT /label= K
FT /note= "a peptide of 4 amino acids"
FT Peptide 45..74
FT /label= B1-B30-Y
FT /note= "human insulin B chain"
FT Peptide 75..86
FT /label= X
FT Peptide 76..86
FT /label= A1-A21-R5
FT /note= "human insulin A chain"
XX
SQ Sequence 86 AA;
XX
Query Match 100.0%; Score 754; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 2.1e-20;
Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREVEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 74

RESULT 2
AAY42866
LC AAY42866 standard; protein; 74 AA.
XX
AC AAY42866;
XX
DT 15-JAN-2000 (first entry)
XX
DE hGH-AAY peptide..hGH-insulin precursor.
XX
KW Insulin precursor; hGH; human growth hormone; insulin A chain;
KW insulin B chain; C-peptide; recombinant production; protein folding;
KW proinsulin; gene.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9967288-A2.
XX
PD 29-DEC-1999.
XX
PF 01-MAR-1999; 99WO-US004422.
XX
PR 01-MAR-1998; 98WO-US003922.
XX
PA (GENE) GENENTECH INC.
XX
CC Claim 2.
XX
DR WPI; 1999-611865/52.
XX
PT New chimeric proteins containing human growth hormone fragments, and
PT particularly for the production of human insulin
XX
PS Claim 2; Page 58; 109pp; English.
XX
CC This sequence represents a chimeric protein. hGH-insulin precursor.
CC This chimeric protein contains an N-terminal fragment of human growth
CC hormone (hGH) of the sequence given in AAY42855, a cleavable peptide
CC linker (AAY42857), and a human insulin precursor comprising insulin
CC A and B chains (AAY42859). The hGH portion of the chimeric protein acts
CC as an intramolecular chaperone (IMC) for the insulin precursor,
CC ensuring it is folded correctly. The cleavable peptide linker has a
CC N-terminal Arg residue which separates the hGH portion of the
CC chimeric protein from the remainder after it has taken part in the
CC formation of recombinant human insulin. Such an hGH-proinsulin chimeric protein
CC provide human insulin with correctly linked cystine bridges with
CC fewer necessary processing steps, and hence results in a higher yield
CC of human insulin. The IMC sequences not only protect insulin sequences
CC from intracellular degradation by a microorganism host, but also promote
CC the folding of the fused insulin precursor, facilitate the purification of
CC the fused protein, and decrease the intermolecular interaction among
CC the fused proteins, thus allowing a higher level of the fused protein to
CC be successfully secreted into culture media. The process steps
CC disclosed include recovery, refolding, enzymatic cleavage and further
CC purification steps, which have to be carried out with the use of low
CC concentrations of mercaptan or the use of hydrophobic absorbent resins.
XX
SQ Sequence 74 AA;

Query Match 100.0%; Score 754; DB 4; Length 74;
Best Local Similarity 100.0%; Pred. No. 2.1e-20;
Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREVEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 74

Db 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREVEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 74

RESULT 8
AAR08627
ID AAR08627 standard; protein; 116 AA.
XX
AC AAR08627;
XX
DT 15-JUN-1997 (first entry)
XX
DE SP1 proinsulin hybrid polypeptide.
XX
KW Insulin; proinsulin; hybrid polypeptide; protein folding;
KW enzymatic cleavage; cyanogen bromide; sulphitolysis.
XX
OS Homo sapiens.
XX
PN WO9600241-A1.
XX
PD 04-JAN-1996.
XX
PF 29-JUN-1994; 94WO-US07268.
XX
PR 29-JUN-1994; 94WO-US07268.
PR 04-JAN-1995; 95GB-0000042.
XX
PA (NOVO) NOVO-NORDISK AS.
XX
PI Sorensen AR, Hansen JB, Markussen J;
XX
DR WPI, 1996-077761/08.
DR N-PSDB; AAI14679.
XX
PT Recombinant insulin produ by correctly folding pro-insulin hybrid
PT polypeptide - then enzymatic cleavage of folded produ t, does not
PT require sulphite protection of SH nor use of cyanogen bromide
XX
PS Example 1B; Figure 7; 66pp; English.
XX
CC A new method for the production of recombinant human insulin
CC comprises folding a hybrid polypeptide comprising proinsulin under
CC conditions that permit correct disulphide bond formation, and
CC subjecting the folded protein to enzymatic cleavage. The insulin
CC produced can then be purified. This sequence is a Ser3 insulin B
CC chain Arg-insulin A chain hybrid polypeptide and is encoded by
CC the plasmid construct pBAS1-IA1. Transformation of the proper
CC E.coli host cells with pBAS1-IA1 results in the efficient
CC expression of the proinsulin hybrid polypeptide, useful for human
CC insulin production. The method produces recombinant human insulin
CC identical to the natural hormone. Hazardous and cumbersome
CC procedures involving cyanogen bromide and sulphitolysis to protect
CC SH groups are avoided since the uncut hybrid polypeptide folds
CC efficiently to the native structure even with the leader attached
CC and Cys unprotected.
XX
SQ Sequence 116 AA;

Query Match [illegible]; Score 294; DB 17; Length 116;
Best Local Similarity [illegible]; Pred. No. 2.5e-29;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 90
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 27 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 116

RESULT 9
AAR71694
ID AAR71694 standard; protein; 137 AA.
XX
AC AAR71694;
XX
DT 20-NOV-1997 (first entry)
XX
DE Mutated insulin precursor Arg31, Arg32 N-terminal extended A-chain.
XX
KW Human insulin precursor; Arg31; Arg32; diabetes; insulin analogue;
KW mating factor alpha 1.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Protein 1..137
FT /label= mating factor alpha 1
FT Peptide 87..115
FT /label= B chain
FT Peptide 117..137
FT /label= A chain
XX
PN WO9719951-A.
XX
PD 05-JUN-1997.
XX
PF 19-SEP-1996; 96WO-EP04041.
XX
PR 17-SEP-1996; 96EP-0306041.
PR 22-NOV-1995; 95EP-0203029.
XX
PA (NOVO) NOVO-NORDISK AS.
XX
PI Andersen AS, Balschmidt JB, Brandenburg S, Thassen J;
PI Markussen J;
XX
DR WPI, 1997-331012/30.
DR N-PSDB; AAQ06425.
XX
PT Acylated insulin derivs. which may be present as a zinc co
PT complex - is used to treat diabetes and is rapid acting.
XX
PS Example 5; Page 70; 106pp; English.
XX
CC AAQ06425 encodes AAR71694 mating factor alpha 1 insulin precursor
CC Arg31, Arg32 N-terminal extended A-chain. The present invention
CC insulin derivatives, in the form claimed herein, may be crystallised
CC complexed via disulphide bonds between cysteines A7/B7 and
CC A20/B19. The derivatives, which may be present as a zinc co
CC complex, can be used as a fast acting treatment for diabetes.
XX
SQ Sequence 137 AA;

Query Match [illegible]; Score 294; DB 17; Length 137;
Best Local Similarity [illegible]; Pred. No. 2.5e-29;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 90
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 48 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 137

RESULT 10
AAR71694
ID AAR71694 standard; protein; 137 AA.
XX
AC AAR71694;
XX
DT 20-NOV-1997 (first entry)
XX
DE Mating factor alpha 1 insulin precursor Arg31, Arg32 N-term. ...
XX
KW Human insulin precursor; Arg31; Arg32; diabetes; insulin analogue;
KW mating factor alpha 1; N-terminal extended A-chain.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Protein 1..137
FT /label= mating factor alpha 1
FT Peptide 87..115

```
RT                     /label= N-terminal peptide
FT   Peptide           95..124
XX                     /label= B-chain
FT   Peptide           125..145
XX                     /label= A-chain
XX
PN   WO9604306-A.
XX
PD   15-FEB-1996.
XX
PF   10-SEP-1994;   94WO-DK00347.
XX
PR   17-SEP-1993;   93DK-00010411.
PR   02-FEB-1994;   94DK-00000082.
XX
PA   (NOVO ) NOVO NORDISK AS.
XX
PI   Andersen AS, Halstrom JB, Hartelund S, Jonasson I;
PI   Markussen J;

DR   WPI; 1996-131811/17.
DR   N-PSDB; AAQ80429.
XX
PT   Acylated insulin derivs. which may be present as a zinc ion
PT   complex - is used to treat diabetes and is rapid acting.
XX
PS   Example 5; Page 85; 140pp; English.
XX
CC   AAQ80429 encodes AAK71694 mating factor alpha 1-insulin precursor
CC   Arg0b1, Arg1b31 N-terminal EEAEAEAER. The insulin precursor comprises
CC   the B and A chains of a claimed human insulin derivative preceded
CC   by the N-terminal amino acids EEAEAEAEr. In the final claimed compsn.
CC   they are covalently connected via disulphide bonds between Cys
CC   residues A7/B7 and A20/B19. The derivative, which may be present
CC   as a zinc ion complex, can be used as a fast action treatment for
CC   diabetes.
XX
SQ   Sequence   145 AA;

  Query Match          100.0%;  Score 294;  DB 16;  Length 145;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-26;
  Matches  52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 EVNQDLISSHLVEALYLVCGERGFFYTPKTRGIVEQCCTSICSLYQLENYCN 52
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94 EVNQDLISSHLVEALYLVCGERGFFYTPKTRGIVEQCCTSICSLYQLENYCN 145


RESULT 14
AAQ71694
ID   AAQ71694 standard; protein; 145 AA.
XX
AC   AAQ71694;
XX
DT   15-JAN-2000  (first entry)
XX
DE   Mating factor alpha 1 insulin precursor Arg0b1, Arg1b31 N-terminal.
XX
KW   Human insulin precursor Arg0b1, Arg1b31; diabetes; zinc ion complex;
KW   mating factor alpha 1; N-terminal EEAEAEAER.
XX
OS   Homo sapiens.
XX
RN   [1]
RP   1-145.
RX   DOI; 10.1016/S0378-1119(99)00335-5.
RX   PUBMED; 10395777.
XX
RA   Andersen A.S., Halstrom J.B., Hartelund S., Jonasson I.;
RT   "Preproinsulin signal peptide alters and delays folding and
RT   secretion of insulin in the yeast Saccharomyces cerevisiae";
RL   Yeast 15:1353-1361(1999).
XX
RN   [2]
RP   1-145.
RA   Andersen A.S., Halstrom J.B., Hartelund S., Jonasson I.;
RT   "Preproinsulin signal peptide alters and delays folding and
RT   secretion of insulin in the yeast Saccharomyces cerevisiae";
RL   Submitted (04-JAN-2000) to the EMBL/GenBank/DDBJ databases.
RL   Novo Nordisk A/S, Novo Alle, DK-2880 Bagsvaerd, Denmark
XX
RN   [3]
RP   1-145.
RA   (NOVO ) NOVO NORDISK AS.
XX
PI   Andersen AS, Halstrom JB, Hartelund S, Jonasson I;
PI   Markussen J;

DR   WPI; 1996-131811/17.
DR   N-PSDB; AAQ80429.
XX
PT   Acylated insulin derivs. which may be present as a zinc ion
PT   complex - is used to treat diabetes and is rapid acting.
XX
PS   Example 5; Page 85; 140pp; English.
XX
CC   AAQ80429 encodes AAK71694 mating factor alpha 1-insulin precursor
CC   Arg0b1, Arg1b31 N-terminal EEAEAEAER. The insulin precursor comprises
CC   the B and A chains of a claimed human insulin derivative preceded
CC   by the N-terminal amino acids EEAEAEAEr. In the final claimed compsn.
CC   they are covalently connected via disulphide bonds between Cys
CC   residues A7/B7 and A20/B19. The derivative, which may be present
CC   as a zinc ion complex, can be used as a fast action treatment for
CC   diabetes.
XX
SQ   Sequence   145 AA;

  Query Match          100.0%;  Score 294;  DB 16;  Length 145;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-26;
  Matches  52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 EVNQDLISSHLVEALYLVCGERGFFYTPKTRGIVEQCCTSICSLYQLENYCN 52
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94 EVNQDLISSHLVEALYLVCGERGFFYTPKTRGIVEQCCTSICSLYQLENYCN 145
```

Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible]

[illegible]

RESULT 15
AAL[illegible]
ID AAL[illegible] standard; protein; [illegible] AA.
XX
AC AAL[illegible];
XX
[illegible] [illegible] (first entry)
XX
DE [illegible] Ala (Bol v2)-Human insulin amorphous, monospherical defi..
XX
KW Bolus; insulin; recombinant preparation; amorphous; monospherical; [illegible]
KW diabetes mellitus.
XX
OS Homo sapiens (produced recombinantly in Escherichia coli).
XX
RN [illegible] A.
XX
RL [illegible] Nov [illegible].
XX
RL [illegible] Apr [illegible].
XX
RL [illegible] Apr [illegible].
XX
RA (IAMH) HOECHST AG.
XX
RT [illegible]; Nielsen K; Obermeier K; Sabel W;
XX
DR WPI; [illegible]/42.
XX
CC Aqueous, monospherical form of insulin derivs. [illegible]
CC diabetes mellitus, are prepared by dissolving in [illegible]
CC isopropanol, are stable when dried or in suspension
XX
FS Example 2; Page 5; 1 pp; German.
XX
CC This sequence is a specific example of an insulin derivative which
CC can be obtained in amorphous, monospherical form by dissolving in
CC an n-propanol/buffer mixture (pH 4.5-6.5) having n-propanol content
CC [illegible]% relative to water. The solution is then diluted with water to
CC [illegible]% n-propanol content at pH [illegible]. The resulting insulin
CC preparation is stable and can be used for the treatment of diabetes
CC mellitus.
XX
SQ Sequence [illegible] AA;

Query Match [illegible]%; Score [illegible]; DB [illegible]; Length [illegible];
Best Local Similarity [illegible]%; Pred. No. [illegible];
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

[illegible]

Search completed: February [illegible], [illegible], 12:44:[illegible]
Job time : [illegible] secs

[illegible]
[illegible] (c) [illegible]

[illegible]

[illegible] [illegible] 12:42:21 ; Search time [illegible] Seconds
(without alignments)
[illegible] Million cell updates/sec

Title: US-09-423-100-5
Perfect score: 2?4
Sequence: 1 FVNQHL[illegible]SHLVEALYLV[illegible]..........[illegible]VEQ[illegible]TSI[illegible]SLYQLENY[illegible]N 52

Scoring table: [illegible]
[illegible]

Searched: 28?224 seqs, [illegible] residues

[illegible] number of hits satisfying chosen parameters: [illegible]

Minimum DB seq length: [illegible]
Maximum DB seq length: [illegible]

Post-processing: Minimum Match [illegible]
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	[illegible]	[illegible]	51	1	INWHP	insulin - sperm wh
2	[illegible]	[illegible]	51	1	INWHF	insulin - finback
3	[illegible]	[illegible]	51	1	INEL	insulin - elephant
4	[illegible]	[illegible]	9?	2	[illegible]82	epidermal growth f
5	[illegible]	[illegible]	51	1	INHY	insulin - hamster
6	[illegible]	[illegible]	51	1	INMSSI	insulin - Egyptian
7	[illegible]	[illegible]	51	2	A[illegible]	insulin precursor
8	[illegible]	[illegible]	11?	1	[illegible]	insulin precursor
9	[illegible]	[illegible]	11?	2	[illegible]	insulin precursor
10	267	[illegible]	110	2	A4217?	insulin precursor
11	267	[illegible]	110	2	[illegible]178	insulin precursor
12	[illegible]	[illegible]	51	1	INWHIS	insulin - sei whal
13	[illegible]	[illegible]	51	1	INGT	insulin - goat
14	[illegible]	[illegible]	51	1	INCMA	insulin - Arabian
15	[illegible]	[illegible]	84	1	[illegible]	insulin precursor
16	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	insulin precursor
17	[illegible]	[illegible]	51	[illegible]	[illegible]	insulin - rat
18	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	insulin precursor
19	[illegible]	[illegible]	51	1	INMKS[illegible]	insulin - common s
20	[illegible]	[illegible].4	11?	2	148[illegible]	insulin precursor
21	[illegible]	[illegible]	10?	1	[illegible]	insulin precursor
22	[illegible]	[illegible].2	51	2	[illegible]	insulin - North Am
23	[illegible]	[illegible]	[illegible]	1	INSH	insulin precursor
24	[illegible]	[illegible]	8?	1	[illegible]	insulin precursor
25	[illegible]	[illegible]	51	1	[illegible]	insulin - Chinchil
26	[illegible]	[illegible].4	1??	2	A[illegible]	insulin precursor
27	[illegible]	[illegible]	1??	[illegible]	INMS[illegible]	insulin I precurso
28	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	insulin I precurso
29	[illegible]	[illegible]	51	1	[illegible]	insulin - goose
30–45	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

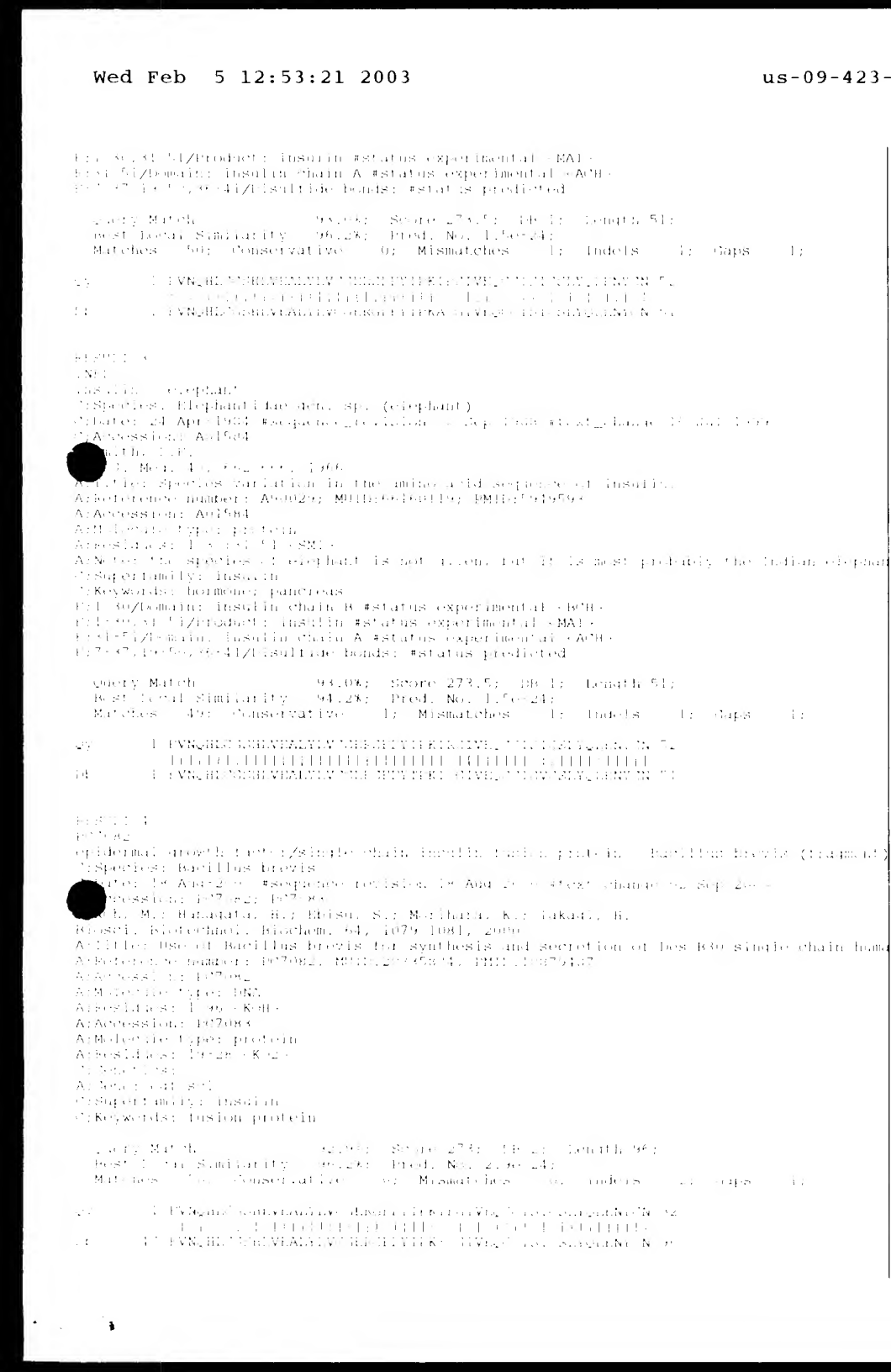
ALIGNMENTS

RESULT 1
INWHP
insulin - sperm whale
C;Species: Physeter [illegible] (sperm whale)
C;Date: [illegible]
C;Accession: A[illegible]; A[illegible]
R;Ishihara, Y.; Saito, T.; [illegible]
Nature [illegible]
A;Title: [illegible]
A;Reference number: A[illegible]
A;Accession: A[illegible]
A;Molecule type: protein
A;Residues: [illegible]
R;Harris, [illegible]; Sanger, [illegible]
Arch. Biochem. Biophys. [illegible]
A;Title: Species differences in insulin.
A;Reference number: A[illegible]
A;Accession: A[illegible]
A;Molecule type: protein
A;Residues: [illegible]
C;Superfamily: insulin
C;Keywords: [illegible]
F;[illegible]/Domain: insulin [illegible] #status experimental [illegible]
F;[illegible]/Domain: insulin [illegible] #status experimental [illegible]
F;[illegible]/Domain: insulin chain A #status experimental [illegible]
F;7-47,19-5?,6-11/Disulfide bonds: #status predicted

Query Match [illegible]
Best Local Similarity [illegible]
Matches [illegible] Mismatches [illegible]

Qy 1 FVNQHL[illegible]SHLVEALYLV[illegible]
Db 1 FVNQHL[illegible]SHLVEALYLV[illegible]GFFYTPKA [illegible]

RESULT 2
INWHF
insulin - finback whale [illegible]
C;Species: [illegible]
C;Date: [illegible]
C;Accession: A[illegible]
R;Hama, H.; [illegible]
J. Biochem. [illegible] 1964
A;Title: The amino acid sequence of [illegible]
A;Reference number: A[illegible]
A;Accession: A[illegible]
A;Molecule type: protein
A;Residues: [illegible]
C;Superfamily: insulin
C;Keywords: [illegible]
F;[illegible]/Domain: insulin chain B #status experimental [illegible]



RESULT 5
INHY
insulin - human
C;Species: Homo sapiens (man)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-1999
C;Accession: A91196
R;Neelon, F.A.; Belcher, H.K.; Steinman, H.; Lebovitz, H.E.
Fed. Proc. 32, 800, 1973
A;Title: Purification and characterization of human insulin A-chain by high-pressure liquid chromatography.
A;Reference number: A91195
A;Accession: A91195
A;Molecule type: protein
A;Residues: 1-30;31-51 <NEE>
C;Superfamily: insulin
C;Keywords: hormone; pancreas
F;1-30/Domain: insulin chain B #status experimental <B2H>
F;31-51/Product: insulin #status experimental <MAT>
F;31-51/Domain: insulin chain A #status experimental <A2H>
F;7-47,12-65,20-41/disulfide bonds: #status predicted

Query Match 51.0%; Score 276; DB 1; Length 51;
Best Local Similarity 94.2%; Pred. No. 2.6e-24;
Matches 49; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 90
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 90

RESULT 6
INMSSP
insulin - Egyptian spiny mouse (fragments)
C;Species: Acomys cahirinus (Egyptian spiny mouse)
C;Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 09-Jul-1999
C;Accession: A01591
R;Buenzli, H.F.; Humbel, R.E.
Hoppe-Seyler's Z. Physiol. Chem. 353, 414-450, 1972
A;Title: Isolation and partial structural analysis of insulin from Acomys cahirinus (Egyptian spiny mouse).
A;Reference number: A01591; MUID:73159071; PMID:4624160
A;Contents: composition
A;Accession: A01591
A;Molecule type: protein
A;Residues: 1-<51> <BUE>
C;Superfamily: insulin
C;Keywords: hormone; pancreas
F;1-29/Domain: insulin chain B #status predicted <B2H>
F;31-51/Product: insulin #status predicted <MAT>
F;31-51/Domain: insulin chain A #status predicted <A2H>
F;7-47,12-65,20-41/disulfide bonds: #status predicted

Query Match 51.0%; Score 276; DB 1; Length 51;
Best Local Similarity 92.9%; Pred. No. 5.1e-24;
Matches 49; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 90
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 FVRQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 90

RESULT 7
A91151
insulin precursor - jack bean (fragments)
N;Alternate names: hypoglycemic agent; plant insulin
C;Species: Canavalia ensiformis (jack bean)
C;Date: 27-Jun-1980 #sequence_revision 27-Jun-1980 #text_change 09-Jul-1999
C;Accession: A91151; A91151
R;Yeriswamy, Antonia; Mahmood, Tahira; Qamar, Abdul Moiz; Ahmad, Nafees; Ali, Syed Asif; Ali, Syed Nawab
Protein Pept. Lett. 6, 15-21, 1999
A;Title: Jack bean seed coat contains a protein with complete sequence homology to the
A;Reference number: A91151
A;Accession: A91151
A;Molecule type: protein
A;Residues: 1-30 <MAC>

A;Reference number: A42179; MUID:92219953; PMID:1569757
A;Accession: A42179
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-110 <SEI>
A;Cross-references: EMBL:X61029; NID:g38251; PIDN:CAA43411.1; PID:g38252
A;Note: sequence extracted from NCBI backbone (NCBIN:25007)
C;Genetics:
A;Introns: 63/3
C;Superfamily: insulin

Query Match 90.8%; Score 267; DB 2; Length 110;
Best Local Similarity 90.5%; Pred. No. 1.5e-23;
Matches 52; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKT----- 30
||||||||||||||||||||||||||||||
Db 25 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 84

Qy 31 ----RGIVEQCCTSICSLYQLENYCN 52
||||||||||||||||||||||
Db 85 SLQKRGIVEQCCTSICSLYQLENYCN 110

RESULT 10
A42179
insulin precursor - chimpanzee
C;Species: Pan troglodytes (chimpanzee)
C;Date: 04-Mar-1993 #sequence_revision 10-N... #text_change 15-Jul-1999
C;Accession: A42179; S22058
R;Soumo, S.; Bell, G.I.; Li, W.H.
Mol. Biol. Evol. 9, 193-203, 1992
A;Title: Sequences of primate insulin genes support the hypothesis of a slower rate of
A;Reference number: A42179; MUID:92219953; PMID:1569757
A;Accession: A42179
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-110 <SEI>
A;Cross-references: EMBL:X61029; NID:g38251; PIDN:CAA43411.1; PID:g38252
A;Note: sequence extracted from NCBI backbone (NCBIN:25007)
C;Genetics:
A;Introns: 63/3
C;Superfamily: insulin

Query Match 90.8%; Score 267; DB 2; Length 110;
Best Local Similarity 90.5%; Pred. No. 1.5e-23;
Matches 52; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKT----- 30
||||||||||||||||||||||||||||||
Db 25 FVNQHLCGSHLVEALYLVCGERGFFYTPKTREAEDLQVGQVELGGGPGAGSLQPLALEG 84

Qy 31 ----RGIVEQCCTSICSLYQLENYCN 52
||||||||||||||||||||||
Db 85 SLQKRGIVEQCCTSICSLYQLENYCN 110

RESULT 11
JQ0178
insulin precursor - crab-eating macaque
C;Species: Macaca fascicularis (crab-eating macaque)
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 16-Jul-1999
C;Accession: JQ0178
R;Wetekam, W.; Steinbeck, J.; Hofschneider, P.H.; Mueller-Hill, B.; Montenarh, M.; Winterhoff, H.
Gene 10, 179-184, 1980
A;Title: The nucleotide sequence of cDNA coding for preproinsulin from the primate Ma
A;Reference number: JQ0178; MUID:80288171; PMID:6161262
A;Accession: JQ0178
A;Molecule type: mRNA
A;Residues: 1-110 <WET>
A;Cross-references: GB:J00136; NID:g182721; PIDN:AAA36819.1; PID:g182...
C;Superfamily: insulin
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-54,90-110/Product: insulin #status predicted <MAT>
F;25-54/Domain: insulin chain B #status predicted <BCH>
F;55-89/Domain: insulin connecting C peptide #status predicted <CPT>
F;90-110/Domain: insulin chain A #status predicted <ACH>
F;31-96,43-109,95-100/Disulfide bonds: #status predicted

Query Match 90.8%; Score 267; DB 2; Length 110;
Best Local Similarity 90.5%; Pred. No. 1.5e-23;
Matches 52; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKT----- 30
||||||||||||||||||||||||||||||
Db 25 FVNQHLCGSHLVEALYLVCGERGFFYTPKTREAEDLQVGQVELGGGPGAGSLQPLALEG 84

Qy 31 ----RGIVEQCCTSICSLYQLENYCN 52
||||||||||||||||||||||
Db 85 SLQKRGIVEQCCTSICSLYQLENYCN 110

RESULT 12
INWHIS
insulin - sei whale
C;Species: Balaenoptera borealis (sei whale)
C;Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 16-Jul-1999
C;Accession: A01582
R;Ishihara, Y.; Saito, T.; Ito, Y.; Fujino, M.
Nature 181, 1468-1469, 1958
A;Title: Structure of sperm- and Sei-whale insulins and their [illegible] by whale pops
A;Reference number: A...242
A;Accession: A01582
A;Molecule type: protein
A;Residues: 1-30;31-51 <ISH>
C;Superfamily: insulin
C;Keywords: hormone, pancreas
F;1-30/Domain: insulin chain B #status experimental <BCH>
F;1-30,31-51/Product: insulin #status experimental <MAT>
F;31-51/Domain: insulin chain A #status experimental <ACH>
F;7-37,19-50,36-41/Disulfide bonds: #status predicted

Query Match 89.6%; Score 264.5; DB 1; Length 51;
Best Local Similarity 92.9%; Pred. No. 2.2e-23;
Matches 48; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
||||||||||||||||||||||||||||||| ||||||| | ||||||||||
Db 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKA-RREVEDPQVAQLELGGGLGGLQALALEP 59

RESULT 13
INGT
insulin - goat
C;Species: Capra aegagrus hircus (domestic goat)
C;Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 16-Jul-1999
C;Accession: A01586
R;Smith, L.F.
Am. J. Med. 40, 662-666, 1966
A;Title: Species variation in the amino acid sequence of insulin.
A;Reference number: A98029; MUID:66168119; PMID:5949593
A;Accession: A01586
A;Molecule type: protein

SeqSearch version 5.1.x
Copyright (c) 1993 - 2003 Compugen Ltd.

Protein protein search, using sw model

Run on: February 5, 2003, 12:42:21 ; Search time 4.0 96 Seconds
(without alignments)
462.274 Million cell updates/sec

Title: US-09-423-100-5
Perfect score: 294
Sequence: 1 FVNQHLCGSHLVEALYLVCG.....IVEQCCTSICSLYQLENYCN 52

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476528 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Prob. N.G. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

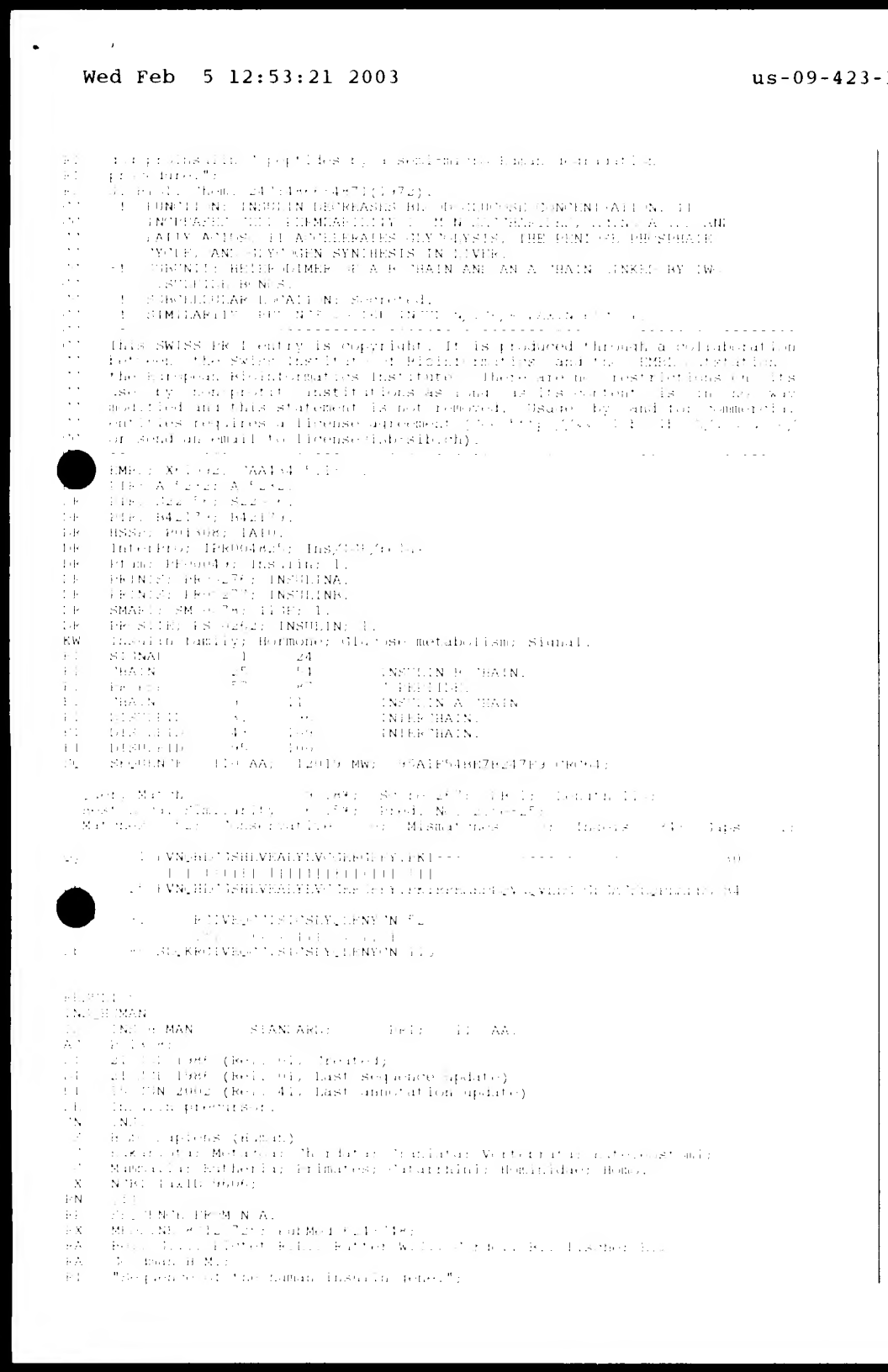
SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	275.5	93.0	51	1	INS_BALPH	P01312 balaenopter
2	272.5	92.7	51	1	INS_PHYMA	P01315 physeter ma
3	268.5	91.3	51	1	INS_ACOCA	P01324 acomys cahi
4	267	90.8	110	1	INS_CERAE	P50487 cercopithec
5	267	90.8	110	1	INS_HUMAN	P01308 homo sapien
6	267	90.8	110	1	INS_MACFA	P50405 macaca fasc
7	267	90.8	110	1	INS_PANTR	P50416 pan troglod
8	264.5	89.6	51	1	INS_BALMU	P01314 balaenopter
9	264.5	89.6	51	1	INS_CAMDR	P01319 camelus dro
10	264.5	89.6	51	1	INS_CAPHI	P01318 capra hircu
11	264	89.5	110	1	INS_PIG	P01315 sus scrofa
12	264	89.5	110	1	INS_RABIT	P01311 oryctolagus
13	263.5	89.6	51	1	INS_FELCA	P01317 felis silv
14	262	89.1	110	1	INS_CANFA	P01321 canis famil
15	260	88.4	110	1	INS_CRIGR	P01314 cricetulus
16	258.5	87.9	110	1	INS_BOVIN	P01317 bos taurus
17	257	87.4	110	1	INS_DASAB	Q62597 psammomys o
18	256.5	87.2	51	1	INS_DIDMA	P16169 didelphis m
19	254.5	86.6	105	1	INS_SHEEP	P01318 ovis aries
20	254	86.4	86	1	INS_HORSE	P01315 equus caball
21	251.5	85.5	51	1	INS_CHIBR	P01327 chinchilla
22	251	85.4	108	1	INS_AOTTR	P10604 aotus trivi
23	250	85.0	109	1	INS1_MOUSE	P01325 mus musculu
24	249	84.7	110	1	INS1_RAT	P01322 rattus norv
25	248.5	84.5	51	1	INS_ANSAN	P07454 anser anser
26	248	84.4	110	1	INS2_MOUSE	P01326 mus musculu
27	248	84.4	110	1	INS2_RAT	P01323 rattus norv
28	246	83.7	52	1	INS_ACIGU	P01328 acipenser g
29	244.5	83.2	51	1	INS_HYSCR	P01320 hystrix cri
30	244.5	83.2	51	1	INS_TRIAM	P79507 trichosurus
31	243.5	82.8	104	1	INS_SERPE	P01468 serrasphorus
32	243.5	82.8	51	1	INS_OMNAN	Q07197 omnithorhyn
33	241.5	82.1	107	1	INS_CHICK	P01332 gallus gall
34	238.5	81.1	51	1	INS_ANAPL	P01316 anas platyr
35	238.5	81.1	51	1	INS_ATEBE	P01321 ateles bele
36	238	81.0	52	2	INS_CERSI	P01306 cervus sika
37	238.5	81.1	51	1	INS_CAVPO	P01325 cavia porcl
38	237.5	80.8	51	1	INS_ONCTS	P01329 oncorhynchu
39	236.5	80.4	51	1	INS_XENLA	P01322 xenopus lae
40	236.5	80.4	51	1	INS2_XENLA	P01323 xenopus lae
41	232	78.9	52	1	INS_PLAMA	P14967 platichthys
42	230	78.2	51	1	INS_HALPE	P04851 pseudopleur
43	228	77.6	52	1	INS_SPHPU	P20657 sphyraena p
44	227	77.2	51	1	INS_NOTCA	P01327 notamacropu
45	227	77.2	51	1	INS_CHICA	P19617 chiloscyllu

ALIGNMENTS

RESULT 1
INS_BALPH
ID INS_BALPH STANDARD; PRT; 51 AA.
AC P01312;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Insulin.
GN INS.
OS Balaenoptera physalus (Fin whale) (Sperm whale).
OS Physeter catodon (Sperm whale) (Physeter macrocephalus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Odontoceti; Mysticeti;
OC Balaenopteridae; Balaenoptera.
OX NCBI_TaxID=9770, 9755;
RN [1]
RP PARTIAL SEQUENCE.
RC SPECIES=Physalus;
RA Nama H., Itani K., Saraki S., Narita K.;
RT "The amino acid sequence in fin whale insulin.";
RL J. Biochem. 56:298-299(1964).
RN [2]
RP SEQUENCE.
RC SPECIES=Catodon;
RA Ishibuna T., Fujita T., Yamaguchi K.;
RT "Structure of sperm whale and fin whale insulins.";
RT whale pepsin.";
RL Nature 181:1469-1470(1958).
RN [3]
RP SEQUENCE.
RC SPECIES=Catodon;
RA Harris J.I., Sanger F., Naughton M.A.;
RT "Species differences in insulin.";
RL Arch. Biochem. Biophys. 65:427-438(1956).
CC -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES THE FORMATION OF PROTEINS, AND
CC FACILITATES TRANSMEMBRANE GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -!- SUBUNIT: HEXAMER (3 A-B CHAIN AND AN A-CHAIN LINKED BY
CC DISULFIDE BONDS).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR PIR; A01008; INWHF.
DR PIR; A01014; INWHP.
DR HSSP; P01317; 1ZNI.
DR InterPro; IPR004825; Insulin.
DR PRINTS; PR00276; INSULINA.
DR PRINTS; PR00277; INSULINB.
DR SMART; SM00078; IlGc; 1.
DR PROSITE; PS00262; INSULIN; 1.
KW Insulin family; Hormone; Disulfide bond; Secreted.
FT CHAIN 1 21 INSULIN B CHAIN.
FT NON_CONS 22 23
FT CHAIN 24 51 INSULIN A CHAIN.
FT DISULFID 7 20 INTERCHAIN.
FT DISULFID 19 24 INTERCHAIN.

```
ID   INSULIN.
AC   INS.
OS   Arvicanthis niloticus (Egyptian spiny mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciuromorpha; Muridae; Murinae; Arvicanthis.
OX   NCBI_TaxID=10068;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92219957; PubMed=1560757;
RA   Bucchini D., Jami J.;
RT   "Isolation and partial structural analysis of insulin genes from two rodent species.";
RT   Arvicanthis) and spiny mouse (Acomys cahirinus).";
RL   Hoppe Seyler's Z. Physiol. Chem. 353:1441-1450(1972).
CC   -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC       INCREASES THE UPTAKE OF GLUCOSE AND
CC       FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC       CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC   -!- SUBUNIT: BEING EITHER 2 A AND B CHAIN AND AN A CHAIN LINKED BY TW
CC       DISULFIDE BONDS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR   EMBL; X52930; CAA37064.1; -.
DR   PIR; A37064; INRT.
DR   MGD; MGI:96482; Ins1/Ins2.
DR   PRINTS; PR00276; INSULINA.
DR   PRINTS; PR00277; INSULINB.
DR   SMART; SM00078; IGFc; 1.
DR   PROSITE; PS00262; INSULIN; 1.
KW   Insulin family; Hormone; Glucose metabolism.
FT   CHAIN         1     30       INSULIN B CHAIN.
FT   NON_TER       30    31
FT   CHAIN         31    51       INSULIN A CHAIN.
FT   DISULFID      7     47       INTERCHAIN.
FT   DISULFID      19    56       INTERCHAIN.
FT   DISULFID      43    48       BY SIMILARITY.
SQ   SEQUENCE   51 AA;  5752 MW;  [illegible] CRC64;

  Query Match             96.0%;  Score 274.5;  DB 1;  Length 51;
  Best Local Similarity   96.2%;  Pred. No. 2.1e-26;
  Matches   50;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREVEDPQVEQLELGGSPGDLQTLALEVARQKRGIVDQCCTSICSLYQLENYCN 80
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREVEDPQVEQLELGGSPGDLQTLALEVARQKRGIVDQCCTSICSLYQLENYCN 80

RESULT 2
INS_ERIEM
ID   INS_ERIEM      STANDARD;      PRT;    51 AA.
AC   P01324;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   Insulin.
GN   INS.
OS   Erinaceus europaeus (Western hedgehog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Protostheria; Erinaceidae; Erinaceus.
OX   NCBI_TaxID=9365;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=86112126; PubMed=3941744;
RA   Conlon J.M., Smith D.D., Yoshuantai A., Thim L.;
RT   "Sequences of hedgehog insulins.";
RL   Am. J. Med. Sci. 291:62-66(1986).
CC   -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC       INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC       FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC       CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC   -!- SUBUNIT: BEING EITHER 2 A AND B CHAIN AND AN A CHAIN LINKED BY TW
CC       DISULFIDE BONDS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- MISCELLANEOUS: THE SPECIES P.EUROPAEUS IS N.I.GIVEN, BUT IT IS
CC       MOST PROBABLY THE INDIAN HEDGEHOG (HEMIECHINUS AURITUS).
CC   -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR   EMBL; X03541; CAA27082.1; -.
DR   PIR; A00782; INEL.
DR   HSSP; P01315; 1ZNI.
DR   InterPro; IPR004825; ins/IGF/relax.
DR   PRINTS; PR00276; INSULINA.
DR   PRINTS; PR00277; INSULINB.
DR   SMART; SM00078; IGFc; 1.
DR   PROSITE; PS00262; INSULIN; 1.
KW   Insulin family; Hormone; Glucose metabolism.
FT   CHAIN         1     30       INSULIN B CHAIN.
FT   NON_TER       30    31
FT   CHAIN         31    51       INSULIN A CHAIN.
FT   DISULFID      7     47       INTERCHAIN.
FT   DISULFID      19    56       INTERCHAIN.
FT   DISULFID      43    48       BY SIMILARITY.
SQ   SEQUENCE   51 AA;  5752 MW;  [illegible] CRC64;

  Query Match             94.1%;  Score 268.5;  DB 1;  Length 51;
  Best Local Similarity   94.2%;  Pred. No. 2.1e-26;
  Matches   49;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREVEDPQVEQLELGGSPGDLQTLALEVARQKRGIVDQCCTSICSLYQLENYCN 80
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREVEDPQVEQLELGGSPGDLQTLALEVARQKRGIVDQCCTSICSLYQLENYCN 80

RESULT 3
INS_ANTCA
ID   INS_ANTCA      STANDARD;      PRT;    51 AA.
AC   P01323;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
```



RX MEDLINE; [illegible]; PubMed [illegible];
RA Snider R., Vick A.;
RT "Synthesis of peptides with the properties of human proinsulin C
RT peptides (hC peptide). 3. Synthesis of the sequences 14-17 and 9-13
RT of human proinsulin C peptides.";
RL Int. J. Pept. Protein Res. [illegible].
RN [24]
RP SYNTHESIS OF RT 87.
RX MEDLINE [illegible]; PubMed [illegible];
RA Gattner K., Jaouen G., Kovacs W., Trauth G.;
RT "Synthesis of peptides with the properties of human proinsulin C
RT peptides (hC peptide). 1. Scheme for the synthesis and preparation of
RT the sequence 28-31 of human proinsulin C peptides.";
RL Chem. Ber. [illegible] (1975).
RN [25]
RP VARIANT LOS ANGELES.
RX MEDLINE [illegible]; PubMed [illegible];
RA Haneda M., Chan S.J., Kwok S.C.M., Rubenstein A.H., Steiner D.F.;
RT "Studies on mutant human insulin genes: identification and sequence
RT analysis of a gene encoding [SerB24]insulin.";
RL Proc. Natl. Acad. Sci. U.S.A. [illegible] (1983).
RN [26]
RP VARIANT LOS ANGELES AND CHICAGO.
RX MEDLINE [illegible]; PubMed [illegible];
RA Shoelson S., Plocksa M., Haneda M., Nahum A., Musso G., Kaiser E.T.,
RA Rubenstein A.H., Tager H.;
RT "Identification of a mutant human insulin predicted to contain a
RT single D-phenylalanine substitution.";
RL Proc. Natl. Acad. Sci. U.S.A. [illegible] (1983).
RN [27]
RP VARIANT CHESTER.
RX MEDLINE [illegible]; PubMed [illegible];
RA Shah S.C., Soltz S., Doppelt P.A., Schwartz E., Steiner D.F.;
RT "A mutation in the B chain coding region is associated with impaired
RT proinsulin conversion in a family with hyperproinsulinemia.";
RL Proc. Natl. Acad. Sci. U.S.A. [illegible] (1987).
RN [28]
RP VARIANT WAKAYAMA.
RX MEDLINE [illegible]; PubMed [illegible];
RA Sakura N., Iwama Y., Yukawa S., Kuzuya T., Hirata H.;
RT "Structurally abnormal insulin in a diabetic patient. Characterization
RT of the variant insulin Asp (Val- 1eu) isolated from the pancreas.";
RL J. Clin. Invest. [illegible] (1986).
RN [29]
RP VARIANT HISTIDINE.
RX MEDLINE [illegible]; PubMed [illegible];
RA Kobori O., Eguchi N., Kadowaki T., Sara A., Akanuma Y., Yamada K.,
RA Motoyoshi S.A., Inoue S.I., Kuzuya T.;
RT "Two unrelated patients with familial hyperproinsulinemia due to a
RT mutation substituting histidine for arginine at position 65 in the
RT B-C junction: molecular identification of the mutation by direct
RT sequencing of genomic deoxyribonucleic acid amplified by polymerase
RT chain reaction.";
RL J. Clin. Endocrinol. Metab. [illegible] (1990).
RN [30]
RP VARIANT HISTIDINE.
RX MEDLINE [illegible]; PubMed [illegible];
RA Yano H., Kitano N., Morimoto M., Nakanishi K.S., Imura H., Seino Y.;
RT "A novel point mutation in the human insulin gene giving rise to
RT hyperproinsulinemia (proinsulin Kyoto).";
RL J. Clin. Invest. Med. [illegible] (1992).
RN [31]
RP STRUCTURE BY NMR.
RX MEDLINE [illegible]; PubMed [illegible];
RA Hua Q., Weiss M.A.;
RT "Toward the solution structure of human insulin: sequential 2D 1H NMR
RT assignment of a des-pentapeptide analog and comparison with the
RT structure.";
RL Biochemistry [illegible] (in press).
RN [32]
RP STRUCTURE BY NMR.
RX MEDLINE [illegible]; PubMed [illegible];
RA Hua Q., Weiss M.A.;
RT "Comparative 2D NMR studies of human insulin and des-pentapeptide
RT insulin: sequential resonance assignment and implications for protein
RT dynamics and receptor recognition.";
RL Biochemistry [illegible] (1991).
RN [33]
RP STRUCTURE BY NMR.
RX MEDLINE [illegible], PubMed [illegible];
RA Hua Q., Weiss M.A.;
RT "Two-dimensional NMR studies of des (B26-B30) insulin: sequence-
RT specific resonance assignments and effects of a short chain sequence.";
RL Biochim. Biophys. Acta [illegible] (in press).
RN [34]
RP STRUCTURE BY NMR.
RX MEDLINE [illegible], PubMed [illegible];
RA Hua Q., Weiss M.A.;
RT "Three-dimensional structure of an insulin dimer. A study of
RT the B9(Asp) mutant of human insulin using nuclear magnetic resonance,
RT distance geometry, and restrained molecular dynamics.";
RL J. Mol. Biol. 227:1148 1.63(1992).
RN [35]
RP STRUCTURE BY NMR OF VARIANT LOS ANGELES.
RX MEDLINE [illegible], PubMed [illegible];
RA Hua Q., Shoelson S.E., Inouye K., Weiss M.A.;
RT "Paradoxical structure and stability in a mutant human insulin
RT associated with diabetes mellitus.";
RL Proc. Natl. Acad. Sci. U.S.A. [illegible] (1995).
RN [36]
RP STRUCTURE BY NMR.
RX MEDLINE [illegible]; PubMed [illegible];
RA Chua X., Bentsen A.M., Burke S.J., Smith [illegible];

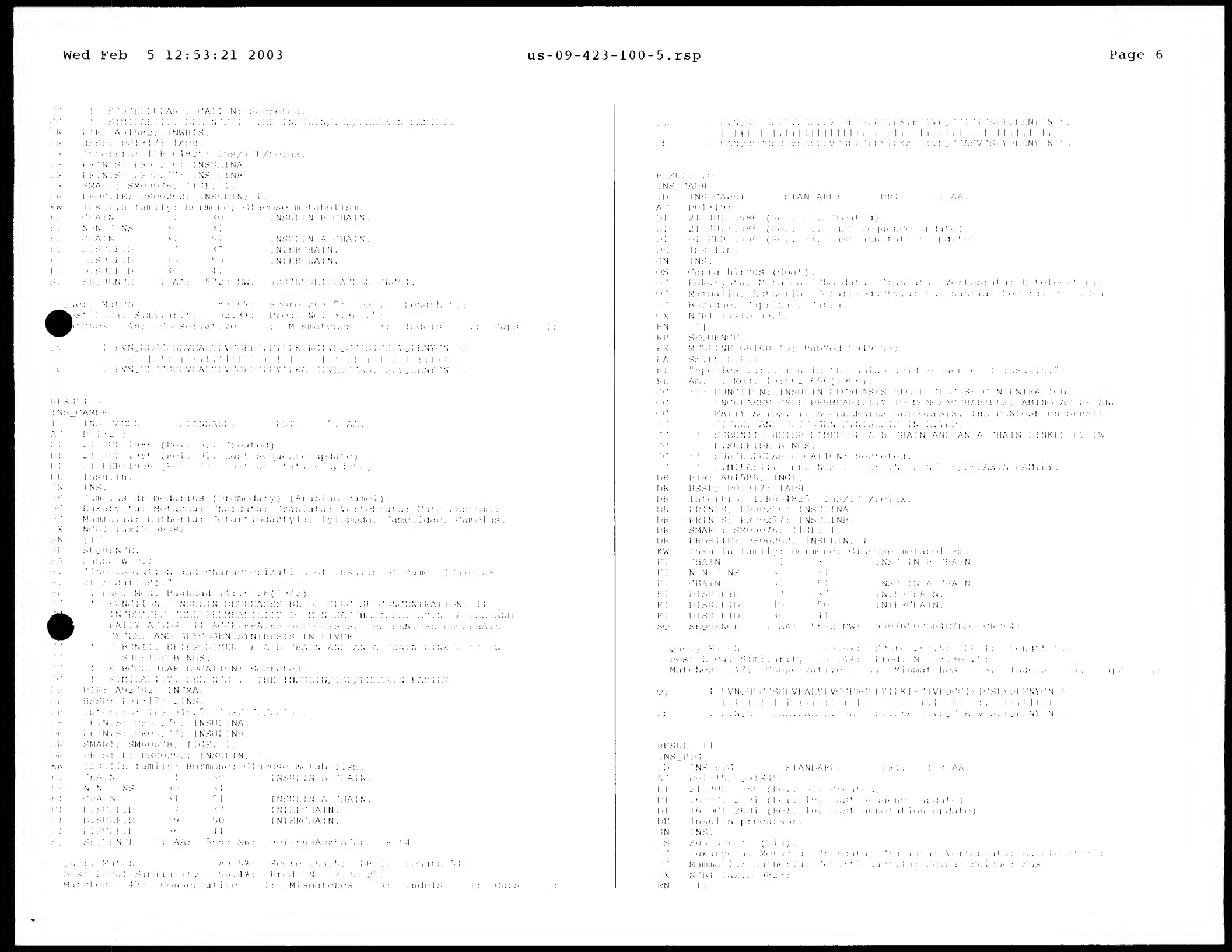
Query Match [illegible]%; Score 264; DB [illegible]; Length [illegible];
Best Local Similarity [illegible]%; Pred. No. 2.5e-25;
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy [illegible] FVNQHLCGSHLVEALYLVCGERGFYTPKT
[illegible]
Db [illegible]

Qy [illegible]
[illegible]
Db [illegible]

RESULT 2
INS_MACFA
ID INS_MACFA STANDARD; PRT; [illegible] AA.
AC P36406; P01309;
DT 21-JUL-1989 (Rel. [illegible], Created)
DT 15-AUG-1997 (Rel. [illegible], Last sequence update)
DT 01-FEB-1999 (Rel. [illegible], Last annotation update)
DE Insulin precursor.
GN INS.
OS Macaca fascicularis (crab-eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE [illegible]; PubMed [illegible];
RA Wittekindt N., Steinbeisser H., Lichtenfels R., Wemmkamper F.,
RA Wittmaack F.;
RT "The nucleotide sequence of cDNA coding for proinsulin of the
RT primate Macaca fascicularis.";
RL Gene [illegible] (1992).

```
CC   * FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC     INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC     FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC     CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC   * SUBUNIT: HEXAMER (ZINC). A DIMER OF TRIMERS. THE TRIMER UNIT MAY
CC     BE MONOMERIC.
CC   * SUBCELLULAR LOCATION: Secreted.
CC   * SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC   -----------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------
DR   EMBL; X00381; AAA31084.1; -.
DR   PIR; A_1234; A_1234.
DR   HSSP; P01308; 2ZNI.
DR   Interpro; IPR004825; Ins/IGF/Relax.
DR   Pfam; PF00049; Insulin; 1.
DR   PRINTS; PR00276; INSULINA.
DR   PRINTS; PR00277; INSULINB.
DR   SMART; SM00078; 1_GF; 1.
DR   PROSITE; PS00262; INSULIN; 1.
KW   Insulin family; Hormone; Glucose metabolism; Signal.
FT   SIGNAL        1     24
FT   CHAIN        25     54       INSULIN B CHAIN.
FT   PROPEP       55     87       C PEPTIDE.
FT   CHAIN        90    110       INSULIN A CHAIN.
FT   DISULFID     31     91       INTERCHAIN.
FT   DISULFID     43    109       INTERCHAIN.
FT   DISULFID     91    103
SQ   SEQUENCE   110 AA;  11862 MW;  B4B1B5A8 A1_RAT CROSSREFERENCE;

  Query Match             ;  Score 267;  DB 1;  Length 110;
  Best Local Similarity   ;  Pred. No. 2.2e-25;
  Matches  52;  Conservative   0;  Mismatches   0;  Indels   34;  Gaps   1;

Qy      1 EVNQLEGGSLQALQKEVAKHKAQPSVAPLAGGLKRQHAMVDFLSAYKCVKQQHGSEVAPC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     27 EVNQLEGGSLQALQKEVAKHKAQPSVAPLAGGLKRQHAMVDFLSAYKCVKQQHGSEVAPC 86

Qy     61 PVNTTLSPEELQLASVEFMMRLKDNARSELMRRFEDVKSRLQVYMGESTESMRGDGTPWAK 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     87 PVNTTLSPEELQLASVEFMMRLKDNARSELMRRFEDVKSRLQVYMGESTESMRGDGTPWAK 146

Qy    121 SLREEDKNLAYLDGFNMLVKQIETMLKFETLEEQSIQNAQDLEKSVLNSHSTYDIGENADH 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    147 SLREEDKNLAYLDGFNMLVKQIETMLKFETLEEQSIQNAQDLEKSVLNSHSTYDIGENADH 206

Qy    181 LSSGDIIKMLESYRDLMVSLGTSKRSLNLYR 211
          |||||||||||||||||||||||||||||||
Db    207 LSSGDIIKMLESYRDLMVSLGTSKRSLNLYR 237

RESULT 2
INS_RAT
ID   INS_RAT        STANDARD;      PRT:   110 AA.
AC   P01308;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-APR-1990 (Rel. 18, Last sequence update)
DT   15-FEB-1995 (Rel. 31, Last annotation update)
DE   Insulin precursor.
GN   INS.
OS   Rattus norvegicus (Norway rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
OC   Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE.
RA   Lemaire S., Gros F., Jacob F., Monod J.;
RL   "Structure of specific messenger RNA coding for the insulin precursor
RL   whale pepsin.";
RL   Nature 301:169-171(1983).
CC   -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC       INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC       FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC       CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC   -!- SUBUNIT: HEXAMER (ZINC). A DIMER OF TRIMERS. THE TRIMER UNIT MAY
CC       BE MONOMERIC.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC   -----------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------
DR   EMBL; X00381; AAA31084.1; -.
DR   PIR; A_1234; A_1234.
DR   HSSP; P01308; 2ZNI.
DR   Interpro; IPR004825; Ins/IGF/Relax.
DR   Pfam; PF00049; Insulin; 1.
DR   PRINTS; PR00276; INSULINA.
DR   PRINTS; PR00277; INSULINB.
DR   SMART; SM00078; 1_GF; 1.
DR   PROSITE; PS00262; INSULIN; 1.
KW   Insulin family; Hormone; Glucose metabolism; Signal.
FT   SIGNAL        1     24
FT   CHAIN        25     54       INSULIN B CHAIN.
FT   PROPEP       55     87       C PEPTIDE.
FT   CHAIN        90    110       INSULIN A CHAIN.
FT   DISULFID     31     91       INTERCHAIN.
FT   DISULFID     43    109       INTERCHAIN.
FT   DISULFID     91    103
SQ   SEQUENCE   110 AA;  11862 MW;  B4B1B5A8 A1_RAT CROSSREFERENCE;

  Query Match             ;  Score 267;  DB 1;  Length 110;
  Best Local Similarity   ;  Pred. No. 2.2e-25;
  Matches  52;  Conservative   0;  Mismatches   0;  Indels   34;  Gaps   1;

Qy      1 EVNQLEGGSLQALQKEVAKHKAQPSVAPLAGGLKRQHAMVDFLSAYKCVKQQHGSEVAPC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     27 EVNQLEGGSLQALQKEVAKHKAQPSVAPLAGGLKRQHAMVDFLSAYKCVKQQHGSEVAPC 86

Qy     61 PVNTTLSPEELQLASVEFMMRLKDNARSELMRRFEDVKSRLQVYMGESTESMRGDGTPWAK 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     87 PVNTTLSPEELQLASVEFMMRLKDNARSELMRRFEDVKSRLQVYMGESTESMRGDGTPWAK 146

Qy    121 SLREEDKNLAYLDGFNMLVKQIETMLKFETLEEQSIQNAQDLEKSVLNSHSTYDIGENADH 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    147 SLREEDKNLAYLDGFNMLVKQIETMLKFETLEEQSIQNAQDLEKSVLNSHSTYDIGENADH 206

Qy    181 LSSGDIIKMLESYRDLMVSLGTSKRSLNLYR 211
          |||||||||||||||||||||||||||||||
Db    207 LSSGDIIKMLESYRDLMVSLGTSKRSLNLYR 237
```



```
RP   SEQUENCE FROM N.A.
RA   Ban X.-Z., Lich R.E.;
RT   "Complete porcine preproinsulin cDNA sequence.";
RL   Submitted (MAY-1990) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 25-109.
RX   MEDLINE=92329187; PubMed=1517256;
RA   Chance R.E., Ellis R.M., Bromer W.W.;
RT   "Porcine proinsulin: characterization and amino acid sequence.";
RL   Science 161:165-167(1968).
RN   [3]
RP   REVISION TO 59.
RA   Chance R.E.;
RT   Submitted (JUL-1970) to the PIR data bank.
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).
RA   Pickens P.I., Dodson G.G., Hodgkin D., Mercola D.;
RT   "Insulin. The structure in the crystal and its reflection in
RT   chemistry and biology.";
RL   Adv. Protein Chem. 26:279-402(1972).
RN   [5]
RP   X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).
RA   Isaacs N.W., Agarwal R.C.;
RT   "Experience with fast Fourier least squares in the refinement of the
RT   crystal structure of rhombohedral 2 zinc insulin at 1.5 A
RT   resolution.";
RL   Acta Crystallogr. A 34:782-791(1978).
RN   [6]
RP   X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).
RX   MEDLINE=99099418; PubMed=9878485;
RA   Baker E.N., Blundell T.L., Cutfield J.F., Cutfield S.M., Dodson E.J.,
RA   Dodson G.G., Crowfoot Hodgkin D.M., Hubbard R.E., Isaacs N.W.,
RA   Reynolds C.D., Sakabe K., Sakabe N., Vijayan N.M.;
RT   "The structure of 2Zn pig insulin crystals at 1.5-A resolution.";
RL   Philos. Trans. R. Soc. Lond., B, Biol. Sci. 319:369-456(1988).
RN   [7]
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX   MEDLINE=92129295; PubMed=1772434;
RA   Ruterjadh P., Hansen F.B., Dodson E., Dodson G., Rotter P.;
RT   "Structure of porcine insulin cocrystallized with clupeine Z.";
RL   Acta Crystallogr. B 47:975-986(1991).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY.
RX   MEDLINE=91282450; PubMed=2025411;
RA   Badger J., Harris M.R., Reynolds C.D., Evans A.C., Dodson E.J.,
RA   Dodson G.G., North A.C.T.;
RT   "Structure of the pig insulin dimer in the cubic crystal.";
RL   Acta Crystallogr. B 47:127-136(1991).
RN   [9]
RP   X-RAY CRYSTALLOGRAPHY (1.65 ANGSTROMS).
RA   Bentley G.A., Wan Z.-L., Chance W.-K., Lipscomb L.A.;
RT   "Structure of monomeric porcine [GlyB1 Bz despentapeptide (B26-B30)
RT   insulin at 1.65 A resolution.";
RL   Acta Crystallogr. D 53:507-512(1997).
CC   -!- FUNCTION: INSULIN STIMULATES UPTAKE OF GLUCOSE BY CONCENTRATING IT
CC       INTRACELLULARLY FOR STORAGE AS GLYCOGEN, FATTY ACIDS, AND
CC       TRIGLYCERIDES. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC       CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC   -!- SUBUNIT: BOTH A AND B CHAINS ARE JOINED TOGETHER BY TWO
CC       DISULFIDE BONDS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC   -!- DATABASE: NAME=Protein Spotlight;
CC       NOTE=Issue 9 - April 2001;
CC       URL="http://www.proteinspotlight.org/";
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
```

```
CC
DR   EMBL; AF004585; AAB[illegible].1; ALT_INIT.
DR   PIR; A01583; IRPG.
DR   PDB; 1ZNI; 20-JAN-94.
DR   PDB; 4INS; 31-JUL-94.
DR   PDB; 6INS; 31-JAN-94.
DR   PDB; 7INS; 31-JAN-94.
DR   PDB; 9INS; 15-[illegible]-95.
DR   PDB; 1IZA; 15-[illegible]-95.
DR   PDB; 1IZB; 15-[illegible]-95.
DR   PDB; 2IFI; 29-JAN-97.
DR   PDB; 1MPJ; 29-JAN-97.
DR   PDB; 3MTH; 29-JAN-97.
DR   PDB; 1LEI; 15-JUN-97.
DR   PDB; 1SDR; 31-AUG-98.
DR   PDB; 1GAV; 28-Feb-97.
DR   PDB; 1EEI; 15-[illegible]-99.
DR   PDB; 1ZNI; 28-JAN-99.
DR   PDB; 1ZNJ; 28-JAN-99.
DR   InterPro; IPR003235; Ins/IGF/relax.
DR   Pfam; PF00049; Insulin; 1.
DR   PRINTS; PR00276; INSULINGENE.
DR   PRINTS; PR00277; INSULINFAM.
DR   SMART; SM00078; IlGF; 1.
DR   PROSITE; PS00262; INSULIN; 1.
KW   Insulin family; Hormone; Glucose metabolism; Signal; Disulfide bond.
FT   SIGNAL        [illegible]    24
FT   CHAIN         25     54       INSULIN B CHAIN.
FT   PROPEP        57     87       C-PEPTIDE.
FT   CHAIN         88    151       INSULIN A CHAIN.
FT   DISULFID      31     83       INTERCHAIN.
FT   DISULFID      91    156       INTERCHAIN.
FT   DISULFID      96    101
FT   HELIX         29     35
FT   STRAND        46     48
FT   HELIX         90     94
FT   HELIX        100    105
FT   STRAND       107    107
SQ   SEQUENCE   110 AA;  11970 MW;  [illegible]2A0EBD [CRC64];

  Query Match             [illegible]% Score 2[illegible]; DB 1; Length [illegible];
  Best Local Similarity   [illegible]%; Pred. No. 3.4e-[illegible];
  Matches   51; Conservative   [illegible]; Mismatches   1; Indels   [illegible]; Gaps   [illegible];

Qy       1 FVNQHLCGSHLVEALYLVCGERGFFYTPKT                              30
           |||||||||||||||||||||||||||||||
Db      25 FVNQHLCGSHLVEALYLVCGERGFFYTPKT                              54

Qy      31 --RGIVEQCCTSICSLYQLENYCN 52
             |||||||||||||||||||||| |
Db      85 QKRGIVEQCCTSICSLYQLENYCN 109
```

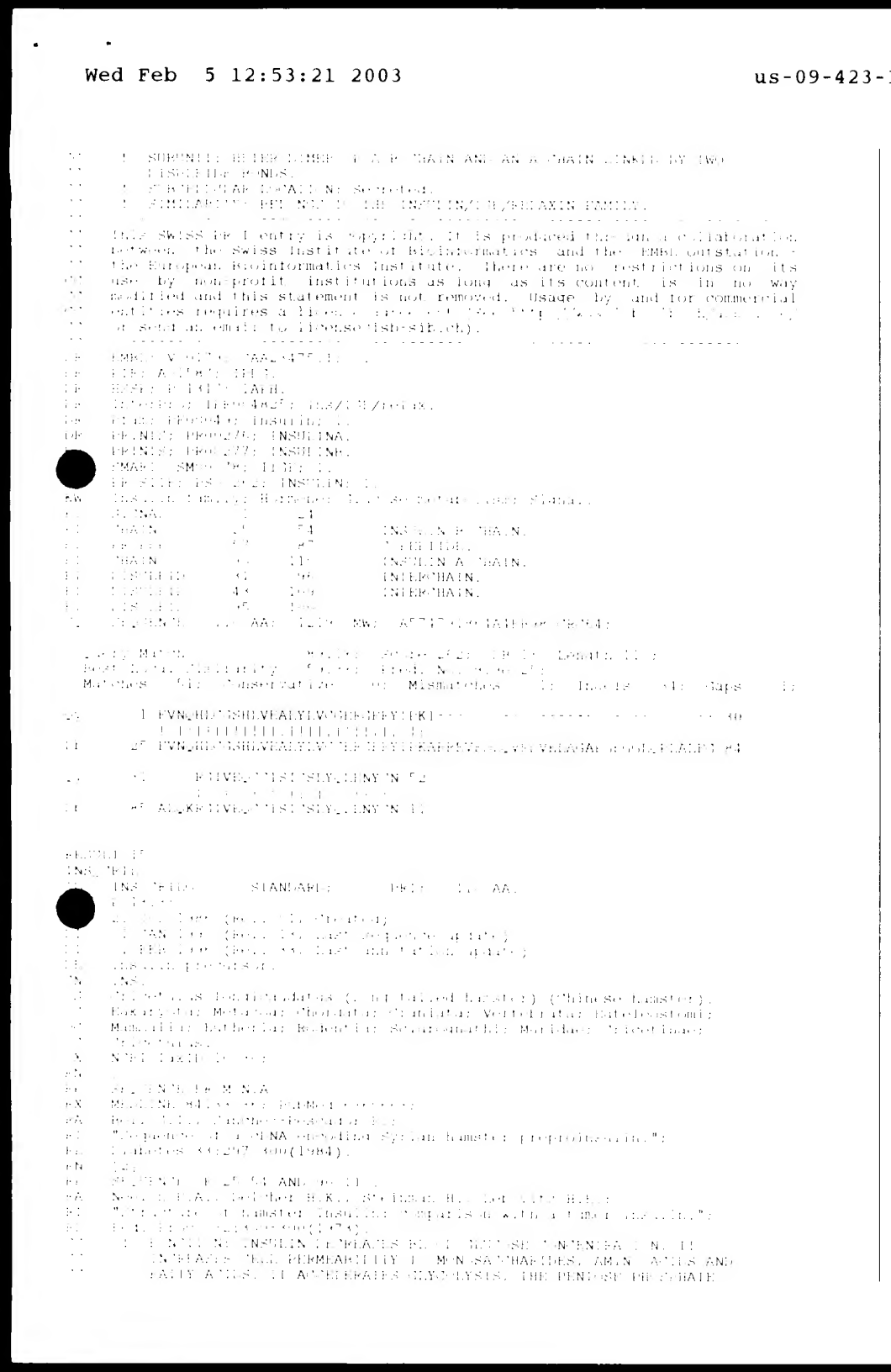
```
RESULT 12
ING_IAB11
ID   [illegible]      STANDARD;      PRT;   110 AA.
AC   P01311;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   Insulin precursor.
GN   INS.
OS   [illegible] [illegible] (rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Laurasiatheria; Artiodactyla; Ruminantia;
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=New Zealand White; TISSUE=Pancreas;
RX   MEDLINE=91179283; PubMed=1672238;
RA   Deshmukh S.D., Gibbins S.D., Balasubramanian T.A., Tattersall [illegible];
RA   Moore W.R., Zahn [illegible];
RT   "Insulin gene expression and [illegible] in mammalian [illegible];
```

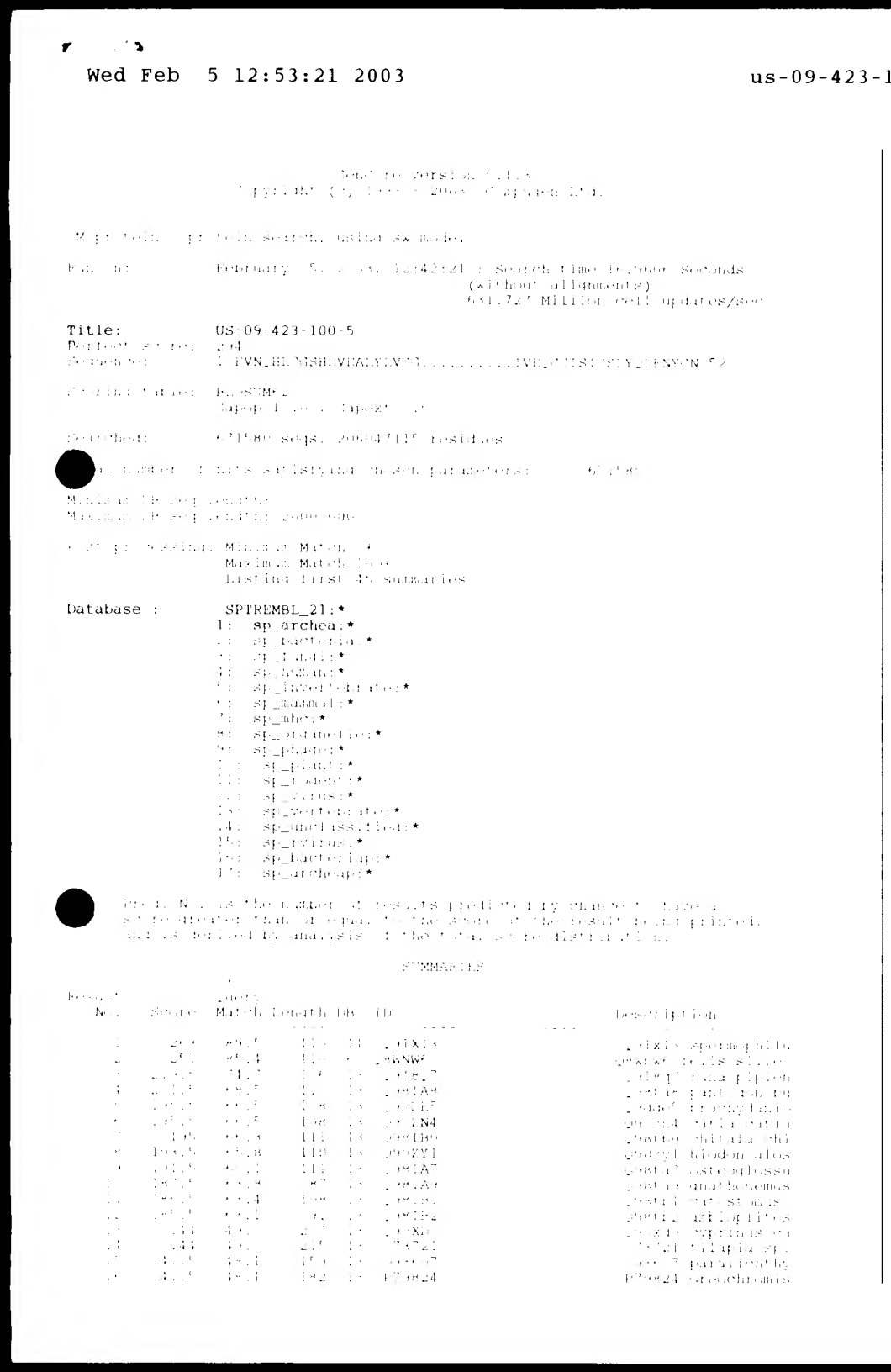
RESULT 14
INS_CANFA
ID INS_CANFA STANDARD; PRT; 51 AA.
AC P01321;
DT 21-JUL-1986 (Rel. 1, Created)
DT 21-JUL-1986 (Rel. 1, Last sequence update)
DT 01-OCT-2001 (Rel. 40, Last annotation update)
DE Insulin.
GN INS.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86218078; PubMed=3518065;
RA Kwok S.C.M., Chan S.J., Steiner D.F.;
RT "Cloning and nucleotide sequence analysis of the dog insulin I and
RT coded amino acid sequence of canine preproinsulin products in
RT additional C peptide fragment.";
RL J. Biol. Chem. 261:7257-7260(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86218079; PubMed=3518066;
RA Smith L.F.;
RT "Species variation in the amino acid sequence of insulin.";
RL Am. J. Med. 40:662-666(1966).
CC -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -!- SUBUNIT: Heterodimer of a B chain and an A chain linked by two
CC disulfide bonds.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the insulin family.
CC ---------- ---------- ---------- ----------
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---------- ---------- ---------- ----------
DR EMBL; M26174; AAA31054.1; -.
DR EMBL; M21158; AAA31254.1; -.
DR PIR; A01581; INRB.
DR HSSP; P01308; 1TYM.
DR InterPro; IPR004825; Ins/IGF/relax.
DR Pfam; PF00049; Insulin; 1.
DR PRINTS; PR00276; INSULINA.
DR PRINTS; PR00277; INSULINB.
DR SMART; SM00078; IGFc; 1.
DR PROSITE; PS00262; INSULIN; 1.
KW Insulin family; Hormone; Glucose metabolism.
FT SIGNAL 1 24
FT CHAIN 25 54 INSULIN B CHAIN.
FT PROPEP 57 87 C PEPTIDE.
FT CHAIN 90 110 INSULIN A CHAIN.
FT DISULFID 31 96 INTERCHAIN.
FT DISULFID 43 109 INTERCHAIN.
FT DISULFID 96 101 BY SIMILARITY.
SQ SEQUENCE 51 AA; 5715 MW; 820A29B8E1D7A4C0 CRC64;

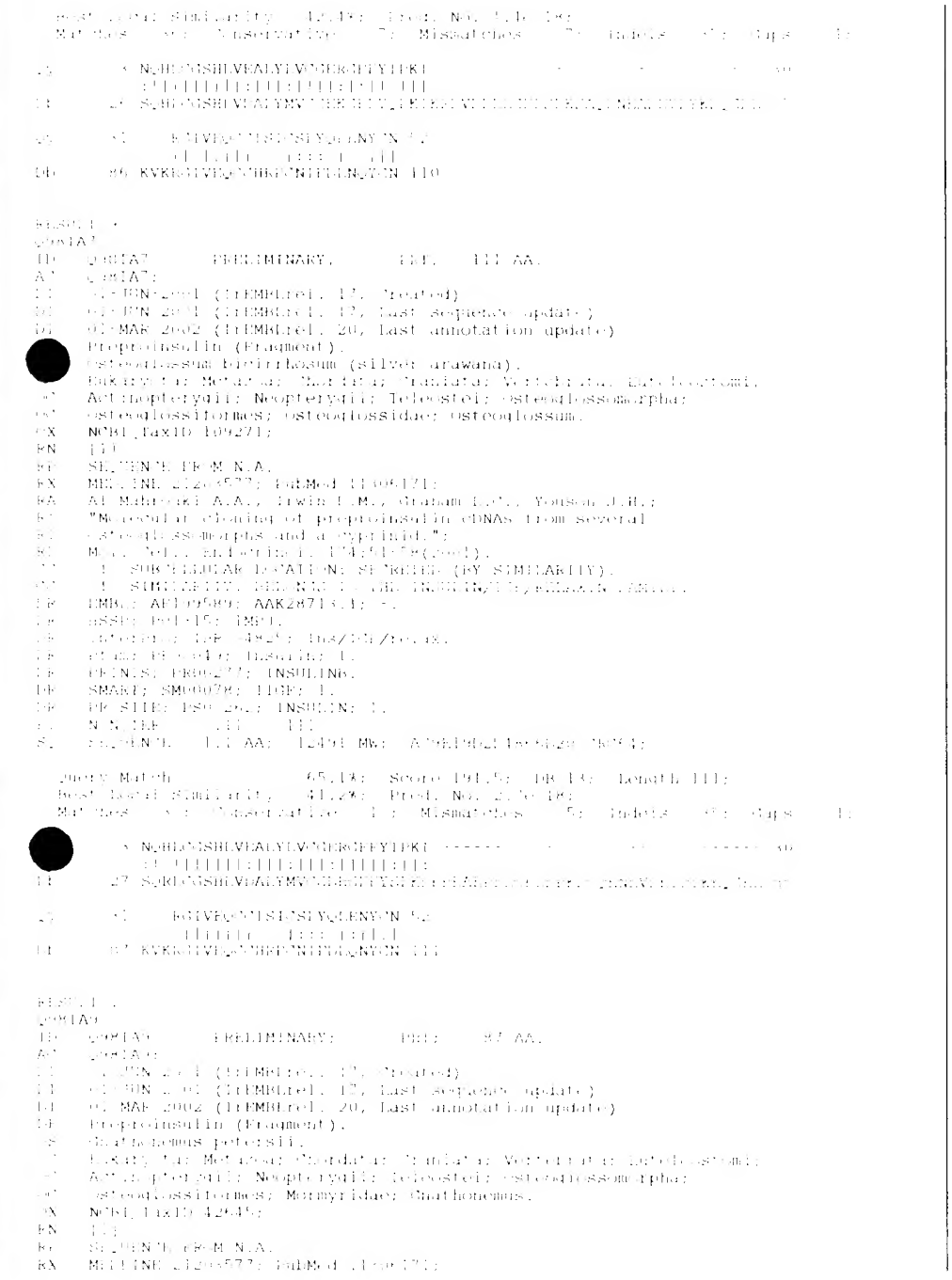
Query Match 100.0%; Score 264; DB 1; Length 110;
Best Local Similarity 100.0%; Pred. No. 2.7e-25;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKT 28
|||||||||||||||||||||||||||||
Db 83 FVNQHLCGSHLVEALYLVCGERGFFYTPKT 110

RESULT 15
INS_CANFA
ID INS_CANFA STANDARD; PRT; 51 AA.
AC P01321;
DT 21-JUL-1986 (Rel. 1, Created)
DT 21-JUL-1986 (Rel. 1, Last sequence update)
DT 01-OCT-2001 (Rel. 40, Last annotation update)
DE Insulin.
GN INS.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86218078; PubMed=3518065;
RA Kwok S.C.M., Chan S.J., Steiner D.F.;
RT "Cloning and nucleotide sequence analysis of the dog insulin I and
RT coded amino acid sequence of canine preproinsulin products in
RT additional C peptide fragment.";
RL J. Biol. Chem. 261:7257-7260(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86218079; PubMed=3518066;
RA Smith L.F.;
RT "Species variation in the amino acid sequence of insulin.";
RL Am. J. Med. 40:662-666(1966).
CC -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.







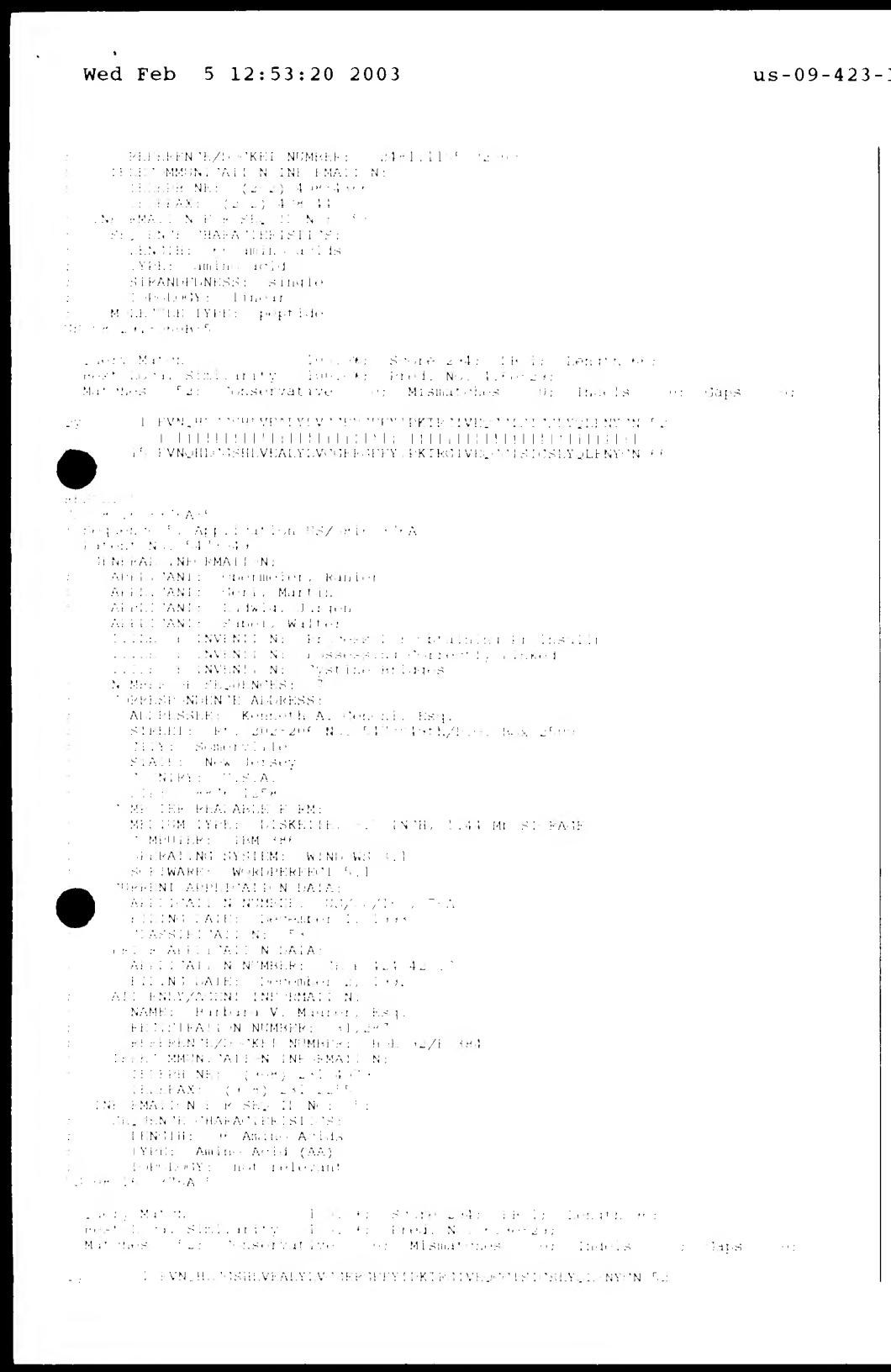
```
RT   in mammalian CHO cell line using green fluorescence protein (GFP)
RT   tagging: secretory production of big IGF1-GFP fusion proteins from
RT   stable transfected CHO cell culture.";
RL   Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN
DR   EMBL; AJ010602; CAA09267.1; -.
DR   HSSP; P01343; 1GF1.
DR   InterPro; IPR004825; Ins/IGF/relax.
DR   Pfam; PF00049; Insulin; 1.
DR   PRINTS; PR00276; INSULINA.
DR   PRINTS; PR00277; INSULINB.
DR   SMART; SM00078; IlGF; 1.
DR   PROSITE; PS00262; INSULIN; 1.
SQ   SEQUENCE   159 AA;  17541 MW;  0B611C8583E0865 CRC64;

  Query Match             48.1%;  Score 141.5;  DB 13;  Length 159;
  Best Local Similarity   47.3%;  Pred. No. 2.2e-11;
  Matches   26;  Conservative    8;  Mismatches   12;  Indels    9;  Gaps    1;

Qy       1 LVGSHLVEALYLVCGERGFFYTPKT---------KREVEGPQVGTLELGGSLYQSLENYC 51
           ||| || |  |||||||::  |          |||::||   | | :|| ||
Db      19 LVGAELVEALYFVCGERGFFYSPKT-GYGSSSRREAEDLQVGQVELGGGPGAGSLQPLALEGSL 77

Search completed: February  5, 2003, 12:47:30
Job time : 17.9606 secs
```

Title: US-09-423-100-5

Database : Issued_Patents_AA:*

SUMMARIES

RESULT 1



```
;       CITY:  New York
;       STATE:  New York
;       COUNTRY:  United States of America
;       ZIP:  10171-6401
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  Patentin Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/423,100
;       FILING DATE:  04-MAR-1995
;       CLASSIFICATION:  514
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Lambiris, Elias J.
;       REGISTRATION NUMBER:  33,728
;       REFERENCE/DOCKET NUMBER:  5985.222-US
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  212-867-0123
;       TELEFAX:  212-878-9655
;   INFORMATION FOR SEQ ID NO:  99:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  137 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-423-100-99

  Query Match             100.0%;  Score 294;  DB 1;  Length 137;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-28;
  Matches  52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 EVNLDELSSELVEALYLVDNDFPEYTIEKTAQIVEQAVTSLSERLYGLENVFN 54
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        86 EVNLDELSSELVEALYLVDNDFPEYTIEKTAQIVEQAVTSLSERLYGLENVFN 139

RESULT 9
US-08-400-256-47
; Sequence 47, Application US/08400256
; Patent No. 5705347
; GENERAL INFORMATION:
;   APPLICANT:  Haviland, Scott
;   APPLICANT:  Halstrom, John
;   APPLICANT:  Jeanssen, IE
;   APPLICANT:  Anderson, Asser Sloth
;   APPLICANT:  Markussen, Jan
;   TITLE OF INVENTION:  ANTIGEN ENGRAIN
;   NUMBER OF SEQUENCES:  49
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  Novo Nordisk of North America, Inc.
;     STREET:  405 Lexington Avenue, 64th Floor
;     CITY:  New York
;     STATE:  New York
;     COUNTRY:  United States of America
;     ZIP:  10174-6401
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  Patentin Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/400,256
;     FILING DATE:  04-MAR-1995
;     CLASSIFICATION:  514
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  Lambiris, Elias J.
;     REGISTRATION NUMBER:  33,728
;     REFERENCE/DOCKET NUMBER:  5985.222-US
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  212-867-0123
;     TELEFAX:  212-878-9655
;   INFORMATION FOR SEQ ID NO:  47:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  145 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-400-256-47

  Query Match             100.0%;  Score 294;  DB 1;  Length 145;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-28;
  Matches  52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 EVNLDELSSELVEALYLVDNDFPEYTIEKTAQIVEQAVTSLSERLYGLENVFN 54
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        94 EVNLDELSSELVEALYLVDNDFPEYTIEKTAQIVEQAVTSLSERLYGLENVFN 147

RESULT 10
US-08-423-100-5
```

; Type: Amino Acid
; Strandedness: Linear
; Molecule Type: protein

Query Match 100.0%; Score 204; DB 5; Length 146;
Best Local Similarity 100.0%; Pred. No. 1.1e-28;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

; Sequence 44, Application US/08/423,741A
; Patent No. 5426036
; GENERAL INFORMATION:
; APPLICANT: Keller, Klaus Peter
; APPLICANT: Bross, Guenther Johannes
; APPLICANT: Ohlmann, Eugen
; APPLICANT: Wallmeier, Helmut
; TITLE OF INVENTION: Processes for the Preparation of Foreign
; TITLE OF INVENTION: Proteins in Streptomycetes
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/423,741A
; FILING DATE: 12-MAR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/192,071
; FILING DATE: 09 MAY 1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/430,922
; FILING DATE: 01 NOV 1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/685,777
; FILING DATE: 29 MAR 1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/877,113,864
; FILING DATE: 05-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 38 37 276.8
; FILING DATE: 08-NOV-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 39 27 449.7
; FILING DATE: 19-AUG-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/423,741A
; FILING DATE: 21-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Klinehaer, Michael K.
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 1243.1033
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000

; TELEFAX: 202 408 4100
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 52 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-423-100-5.rai

Query Match 100.0%; Score 204; DB 5; Length 146;
Best Local Similarity 100.0%; Pred. No. 1.1e-28;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 14
US-08-423,617-4
; Sequence 4, Application US/08/423,617
; Patent No. 5466666
; GENERAL INFORMATION:
; APPLICANT: Obermeier, Rainer
; APPLICANT: Sahl, Walter
; APPLICANT: Seitz, Peter
; APPLICANT: Geisen, Karl
; TITLE OF INVENTION: Amorphous Monospherical Forms of Lys-C
; TITLE OF INVENTION: Derivatives
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/423,617
; FILING DATE: 25-APR-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 14 [illegible]
; FILING DATE: 27-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Carol P. Einaudi
; REGISTRATION NUMBER: 33,220
; REFERENCE/DOCKET NUMBER: 2347.2-US-DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202 408 4400
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 52 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli
US-08-423-617-4

Query Match 100.0%; Score 204; DB 12; Length [illegible];
Best Local Similarity 100.0%; Pred. No. 1.1e-28;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Title: US-09-423-100-5

Perfect score: 294

Sequence: 1 EVN_HLYSGHEVEALYYV[illegible]VEL[illegible]SNYN *L

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 12[illegible] seqs, [illegible] residues

[illegible] of 100 satisfying [illegible] searched

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PCT.pep:*
2: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US09_PUBLISHED.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US09A_NEW_PCT.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US09A_PUBLISHED.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US09B_NEW_PCT.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09B_PUBLISHED.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_NEW_PCT.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09C_PUBLISHED.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US09D_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US09D_NEW_PCT.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US09D_PUBLISHED.pep:*

Score: is the number of results printed. Sorted by score.
Query Match: Percent of query sequence that is matched.
Best Local Similarity: Percent of residues that are similar.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	[illegible]	[illegible]	[illegible]	[illegible]	US-09-[illegible]-[illegible]	Sequence 5, Appl...
2	[illegible]	[illegible]	[illegible]	[illegible]	US-09-[illegible]-[illegible]	Sequence 7, Appl...
3	[illegible]	[illegible]	[illegible]	[illegible]	US-09-[illegible]-[illegible]	Sequence 3, Appl...
4	[illegible]	[illegible]	[illegible]	[illegible]	US-09-[illegible]-[illegible]	Sequence 1, Appl...
5	[illegible]	[illegible]	[illegible]	[illegible]	US-09-[illegible]-[illegible]	Sequence 19, Appl...
6	[illegible]	[illegible]	[illegible]	[illegible]	US-09-[illegible]-[illegible]	Sequence 9, Appl...
7	[illegible]	[illegible]	[illegible]	[illegible]	US-09-[illegible]-[illegible]	Sequence 1, Appl...
8	[illegible]	[illegible]	[illegible]	[illegible]	US-09-[illegible]-[illegible]	Sequence 2, Appl...
9	[illegible]	[illegible]	[illegible]	[illegible]	US-09-[illegible]-[illegible]	Sequence 3, Appl...
10	[illegible]	[illegible]	[illegible]	[illegible]	US-09-[illegible]-[illegible]	Sequence 4, Appl...
11	[illegible]	[illegible]	[illegible]	[illegible]	US-09-[illegible]-[illegible]	Sequence 5, Appl...
12	[illegible]	[illegible]	[illegible]	[illegible]	US-09-[illegible]-[illegible]	Sequence 125, App...
13	[illegible]	[illegible]	[illegible]	[illegible]	US-09-[illegible]-[illegible]	Sequence 1, Appl...
14	[illegible]	[illegible]	[illegible]	[illegible]	US-09-[illegible]-[illegible]	Sequence 63, Appl...
15	[illegible]	[illegible]	[illegible]	[illegible]	US-09-[illegible]-[illegible]	Sequence 12, Appl...
16	[illegible]	[illegible]	[illegible]	[illegible]	US-09-[illegible]-[illegible]	Sequence 12, Appl...
17	[illegible]	[illegible]	[illegible]	[illegible]	US-09-[illegible]-[illegible]	Sequence 12, Appl...
18	[illegible]	[illegible]	[illegible]	[illegible]	US-09-[illegible]-[illegible]	Sequence [illegible], Appl...

ALIGNMENTS

RESULT 1
US-09-[illegible]-[illegible]
; Sequence 5, Application [illegible]
; Patent No. US[illegible]
; GENERAL INFORMATION:
; APPLICANT: [illegible]
; TITLE OF INVENTION: [illegible]
; NUMBER OF SEQUENCES: [illegible]
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: [illegible]
; STREET: [illegible]
; CITY: [illegible]
; STATE: [illegible]
; COUNTRY: USA
; ZIP: [illegible]
; COMPUTER READABLE FORM:
; MEDIUM TYPE: [illegible]
; COMPUTER: [illegible]
; OPERATING SYSTEM: [illegible]
; SOFTWARE: [illegible]
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US[illegible]
; FILING DATE: [illegible]
; CLASSIFICATION: [illegible]
; ATTORNEY/AGENT INFORMATION:
; NAME: [illegible]
; REGISTRATION NUMBER: [illegible]
; REFERENCE/DOCKET NUMBER: [illegible]
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: [illegible]
; TELEFAX: [illegible]
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: [illegible] amino acids
; TYPE: amino acid
; STRANDEDNESS: [illegible]
; TOPOLOGY: [illegible]
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 5: [illegible]

RESULT 2
US-09-423-100-5
; Sequence 5, Application US/09423100
; Patent No. US6200764
; GENERAL INFORMATION:
; APPLICANT: Bowman, Barry
; TITLE OF INVENTION: PROTEIN VARIANTS
; FILE REFERENCE: PI7126P1
; CURRENT APPLICATION NUMBER: US/09/423,100
; CURRENT FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: US/60/077,994
; PRIOR FILING DATE: 2000-03-13
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 5
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-423-100-5

Query Match 61.7%; Score 278.5; DB 4; Length 61;
Best Local Similarity 98.1%; Pred. No. 9.8e-27;
Matches 51; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 3
US-09-423-100-3
; Sequence 3, Application US/09423100
; Patent No. US6200764
; GENERAL INFORMATION:
; APPLICANT: Bowman, Barry
; TITLE OF INVENTION: PROTEIN VARIANTS
; FILE REFERENCE: PI7126P1
; CURRENT APPLICATION NUMBER: US/09/423,100
; CURRENT FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: US/60/077,994
; PRIOR FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 3
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-423-100-3

Query Match 100.0%; Score 451.5; DB 4; Length 107;
Best Local Similarity 100.0%; Pred. No. 1.2e-43;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 4
US-09-423-100-4
; Sequence 4, Application US/09423100
; Patent No. US6200764
; GENERAL INFORMATION:
; APPLICANT: Bowman, Barry
; TITLE OF INVENTION: PROTEIN VARIANTS
; FILE REFERENCE: PI7126P1
; CURRENT APPLICATION NUMBER: US/09/423,100
; CURRENT FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: US/60/077,994
; PRIOR FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 4
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-423-100-4

Query Match 100.0%; Score 451.5; DB 4; Length 107;
Best Local Similarity 100.0%; Pred. No. 1.2e-43;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

; PRIOR APPLICATION NUMBER: JP 2000-174691
; PRIOR FILING DATE: 2000-06-12
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-423-100-5

Query Match 90.8%; Score 267; DB 9; Length 86;
Best Local Similarity 68.5%; Pred. No. 3.7e-25;
Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

Qy 1 PVNQHLCGSHLVEALYLVCGERGFFYTPKT 52
PI IIIIIIIIIIIIIIIIIIIIIIIIIIIIII
Db 1 PVNQHLCGSHLVEALYLVCGERGFFYTPKT 52

Qy 53 KSIVEQCCTSICSLYQLENYCN 92
IIIIIIIIIIIIIIIIIIIII
Db 53 KSIVEQCCTSICSLYQLENYCN 92

RESULT 2
US-10-028-410-2
; Sequence 2, Application US/10028410
; Patent No. US20020160955A1
; GENERAL INFORMATION:
; APPLICANT: Dubaquie, Yves
; APPLICANT: Lowman, Henry
; TITLE OF INVENTION: PROTEIN VARIANTS
; FILE REFERENCE: P1712R1-1
; CURRENT APPLICATION NUMBER: US/10/028,410
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US/60/477,924
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 2
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-028-410-2

Query Match 90.8%; Score 267; DB 9; Length 86;
Best Local Similarity 68.5%; Pred. No. 3.7e-25;
Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

Qy 1 PVNQHLCGSHLVEALYLVCGERGFFYTPKT 52
PI IIIIIIIIIIIIIIIIIIIIIIIIIIIIII
Db 1 PVNQHLCGSHLVEALYLVCGERGFFYTPKT 52

Qy 53 KSIVEQCCTSICSLYQLENYCN 92
IIIIIIIIIIIIIIIIIIIII
Db 53 KSIVEQCCTSICSLYQLENYCN 92

RESULT 3
US-10-054-873-4
; Sequence 4, Application US/10054873
; Patent No. US20020164712A1
; GENERAL INFORMATION:
; APPLICANT: Gan, Zhong Ru
; TITLE OF INVENTION: Chimeric Protein Containing an
; Intramolecular Chaperone-Like Sequence
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/054,873
; FILING DATE: 22-Jan-2002
; CLASSIFICATION: 530/399
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/817,952
; FILING DATE: 31-MAR-1998
; APPLICATION NUMBER: US 09/424,103
; FILING DATE: 11-DEC-2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Wang, Frank J
; REGISTRATION NUMBER: 46,946
; REFERENCE/DOCKET NUMBER: 018282-000300US
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 86 amino acids
; TYPE: amino acid
; STRANDEDNESS: Unknown
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-054-873-4

Query Match 90.8%; Score 267; DB 9; Length 86;
Best Local Similarity 68.5%; Pred. No. 3.7e-25;
Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

Qy 1 PVNQHLCGSHLVEALYLVCGERGFFYTPKT 52
PI IIIIIIIIIIIIIIIIIIIIIIIIIIIIII
Db 1 PVNQHLCGSHLVEALYLVCGERGFFYTPKT 52

Qy 53 KSIVEQCCTSICSLYQLENYCN 92
IIIIIIIIIIIIIIIIIIIII
Db 53 KSIVEQCCTSICSLYQLENYCN 92

RESULT 11
US-09-947-563-4
; Sequence 4, Application US/09947563
; Patent No. US20020156234A1
; GENERAL INFORMATION:
; APPLICANT: Rubrecht, Franz-Josef
; Keller, Reinhold
; TITLE OF INVENTION: Improved process for obtaining
; insulin precursors having correctly bonded cystine bridges
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/947,563
; FILING DATE: 07-Sep-2001
; CLASSIFICATION: [illegible]
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: [illegible]
; FILING DATE: Unknown
; ATTORNEY/AGENT INFORMATION:
; NAME: Leslie McDonell
; REGISTRATION NUMBER: 34,872

; TITLE OF INVENTION: DNAS ENCODING NEW FUSION PEPTIDES AND PROCESSES FOR
; TITLE OF INVENTION: PREPARING USEFUL POLYPEPTIDES THROUGH EXPRESSION OF THE
; TITLE OF INVENTION: DNAS
; FILE REFERENCE: 08210.026
; CURRENT APPLICATION NUMBER: US/09/280,030A
; CURRENT FILING DATE: 1999-03-26
; EARLIER APPLICATION NUMBER: 08/[illegible]/1998
; EARLIER FILING DATE: 1998-03-[illegible]
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 63
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Substituted Is
; OTHER INFORMATION: [illegible] Acid Sequence Set
; SHELF INFORMATION: MSRCP; MSRCP-1 Met-Pro-Insulin
US-09-280-030-63

Query Match 99.8%; Score 267; DB 10; Length 117;
Best Local Similarity 100.0%; Pred. No. 5.6e-25;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EVNQHLCGSHLVEALYLVCGERGFFYTPKT- 52
||||||||||||||||||||||||||||||
Db 66 EVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 125

Qy 53 GSLQKRGIVEQCCTSICSLYQLENYCN 79
|||||||||||||||||||||||||||
Db 126 GSLQKRGIVEQCCTSICSLYQLENYCN 152

Search completed: February 5, 2003, 12:48:48
Job time : 42.595 secs